FILED

99 557 SEP 27 1999

OFFICE OF THE CLERK

No. _____

IN THE

Supreme Court of the United States

SOUTHEASTERN PENNSYLVANIA
TRANSPORTATION AUTHORITY,

Petitioner,

v.

CATHERINE NATSU LANNING, ALTOVISE LOVE,
BELINDA KELLY DODSON, DENISE DOUGHERTY,
AND LYNNE ZIRILLI,
and
UNITED STATES OF AMERICA,

Respondents.

ON PETITION FOR WRIT OF CERTIORARI
TO THE UNITED STATES COURT OF APPEALS
FOR THE THIRD CIRCUIT

PETITION FOR WRIT OF CERTIORARI

SAUL H. KRENZEL, ESQUIRE
Saul H. Krenzel & Associates
The Robinson Building
42 South 15th Street, Suite 800
Philadelphia, PA 19102
Counsel for Petitioner

PACKARD PRESS® / APPELLATE DIVISION, 1617 JFK BOULEVARD, PHILA., PA. 19103 (215) 563-9000

20588

QUESTIONS PRESENTED FOR REVIEW

1. Did the Third Circuit correctly construe pre-*Wards Cove* Supreme Court precedent in this gender disparate impact case when it defined the standard of 'business necessity' to require an employer testing entry level applicants to set its cut-off score at only the "minimum qualifications" for the job, *contrary to every 'business necessity' standard ever articulated by this Court and all other Circuits?*

2. Is the Third Circuit correct in holding that public safety should *not* be considered when assessing the employer's burden in satisfying the 'business necessity' defense, or are the 5th, 6th, 8th, 9th, 10th and 11th Circuits correct in articulating a lighter burden of proof for an employer when public safety is at issue?

3. If the Third Circuit's "minimum qualifications" test is accepted by this Court, are the "minimum qualifications" to be set at the level of performance of those incumbent police officers that have failed to meet the required aerobic capacity standard where:

   a. The police department implemented this aerobic capacity test for new applicants for the specific purpose of better protecting the public by requiring performance levels *above that of its incumbent police force*;

   b. the police department already determined that the incumbent officer level of fitness and performance was inadequate to combat crime, and this determination was the impetus for the implementation of its aerobic capacity test;

   c. the police department cannot fire its incumbent police officers for failing to meet the standard of its new test because of its collective bargaining agreement; and

   d. the District Court found that the police department's new physical abilities test correlates with and predicts better performance of critical safety-sensitive tasks, and better overall performance, although it has a disparate impact.

LIST OF PARTIES

The parties to the proceedings below are summarized as follows. Petitioner, Southeastern Pennsylvania Transportation Authority ("SEPTA"), is a regional mass transit authority and Commonwealth agency, pursuant to 74 Pa. Cons. Stat. Ann. § 1701, *et seq.* Respondents are Catherine Natsu Lanning, Denise Dougherty, Altovise Love, Belinda Kelly Dodson and Lynne Zirilli, individually and on behalf of a class consisting of all 1993 female applicants, 1996 female applicants and future female applicants for employment as SEPTA police officers who have been or will be denied employment by reason of their inability to meet the physical entrance requirement of running 1.5 miles in 12 minutes or less. The United States, through the Department of Justice is an additional Respondent, and filed a complaint in accordance with Section 707 of Title VII, 42 U.S.C. §2000e-6 of the Civil Rights Act of 1991, alleging that SEPTA was engaged in a pattern or practice of employment discrimination against females through the use of its physical abilities pre-employment applicant test for the position of transit police officer.

TABLE OF CONTENTS

	Page
QUESTIONS PRESENTED	(i)
LIST OF PARTIES	(ii)
TABLE OF AUTHORITIES	(v)
OPINIONS BELOW	1
STATEMENT OF JURISDICTION	1
STATUTORY PROVISIONS INVOLVED	1
STATEMENT OF THE CASE	1
BASIS FOR FEDERAL JURISDICTION AND DECISIONS BELOW	11
REASONS FOR GRANTING THE WRIT	12
I. SINCE THE SUPREME COURT HAS NOT EXAMINED THE 'BUSINESS NECESSITY' TEST SINCE *WARDS COVE*, OR AFTER THE ENACTMENT OF THE CIVIL RIGHTS ACT OF 1991, THIS COURT SHOULD NOW CLARIFY THE 'BUSINESS NECESSITY' STANDARD AND SETTLE THE CONFLICT IN THE CIRCUITS	12
II. THE "MINIMUM QUALIFICATIONS" TEST OF 'BUSINESS NECESSITY' CONFLICTS WITH EVERY SUPREME COURT AND CIRCUIT COURT FORMULATION OF 'BUSINESS NECESSITY' UNDER TITLE VII	14

TABLE OF CONTENTS—*(Continued)*

Page

III. THE THIRD CIRCUIT'S REFUSAL TO RECOGNIZE PUBLIC SAFETY AS A CONSIDERATION IN ASSESSING THE 'BUSINESS NECESSITY' OF A TEST CONFLICTS WITH DECISIONS OF THIS COURT, AND WITH SIX OTHER CIRCUIT COURTS AND ADVERSELY IMPACTS THE SAFETY OF MILLIONS OF CITIZENS 17

IV. THE THIRD CIRCUIT'S "MINIMUM QUALIFICATIONS" STANDARD PREVENTS EMPLOYERS FROM UPGRADING THE PERFORMANCE OF THEIR WORKFORCE THROUGH APPLICANT TESTING AND SETS ENTRY LEVEL JOB PERFORMANCE AT THE LOWEST COMMON DENOMINATOR FOR THOUSANDS OF EMPLOYERS 20

CONCLUSION..... 25

APPENDIX A – Opinion and Judgment of the Court of Appeals 1a

APPENDIX B – Memorandum Decision of the District Court 46a

APPENDIX C – Statutes Involved 163a

APPENDIX D – Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. §1600, *et seq.* 166a

APPENDIX E – Principles for the Validation and Use of Personnel Selection Procedures (1987)..... 168a

APPENDIX F – Letter filing one copy of the video tape of the 1.5 mile running test given to applicants for the transit officer position in 1993 ..... 171a

TABLE OF AUTHORITIES

Cases:

Page

Albemarle Paper Co. v. Moody, 422 U.S. 405 (1975) 15, 16, 20

Allen, et al v. Energy Corporation, et al., 181 F.3d 902 (8th Cir. 1999) 16

The Association of Mexican-American Educators, 183 F.3d 1055 (9th Cir. 1999) 17

Chrisner v. Complex Auto Transit, Inc., 645 F.2d 1251 (6th Cir. 1981) 18

Cleghorn v. Herrington, 813 F.2d 992 (9th Cir. 1987) 21

Contreras v. City of Los Angeles, 656 F.2d 1267 (9th Cir. 1981) 16

Davis v. City of Dallas, 777 F.2d 205 (5th Cir. 1985)..... 18

Dothard v. Rawlinson, 433 U.S. 321 (1977)..... 13, 16

Fitzpatrick v. City of Atlanta, 2 F.3d 1112 (11th Cir. 1993) 17

Guardians Association of New York City Police Department v. Civil Service Commission, 630 F.2d 79 (2d Cir. 1980), *cert denied*, 452 U.S. 940 (1981)..... 20, 21

Griggs v. Duke Power Co., 401 U.S. 424 (1971)..... 15, 16

Harless v. Duck, 619 F.2d 611 (6th Cir. 1980)..... 16

Harris v. Pan American World Airways, Inc., 649 F.2d 670 (9th Cir. 1980) 19

Hodgson v. Greyhound Lines, Inc., 499 F.2d 859 (7th Cir. 1974) 19

Lanning, et al. v. SEPTA, 181 F.3d 478 (3rd Cir. 1999)..... 1, 8, 13, 17, 22, 23

Lanning, et al. v. SEPTA, 1998 U.S. Dist. LEXIS 9388 (E.D. Pa. 1998) 1

TABLE OF AUTHORITIES—*(Continued)*

Cases: Page

NAACP v. Town of East Haven, 70 F.3d 219 (2d Cir. 1995) 16

New York City Transit Authority v. Beazer, 440 U.S. 568 (1979) 13, 21

Pettway v. American Cast Iron Pipe Co., 494 F.2d 211 (5th Cir. 1974), *cert. denied*, 393 U.S. 1115 (1979)... 16

Police Officers v. City of Columbus, 916 F.2d 1092 (6th Cir. 1990) 21

Robinson v. Lorillard Corporation, 444 F.2d 791 (4th Cir. 1971), *writ of cert. dismissed*, 404 U.S. 1006 (1971).. 16

Smith v. City of DeMoines, 99 F.3d 1466 (8th Cir. 1996) . 16

Spurlock v. United Airlines, Inc., 475 F.2d 216 (10th Cir. 1972) 19

Wards Cove Packing Co. v. Atonio, 490 U.S. 642 (1989) . 11

Washington v. Davis, 426 U.S. 229 (1976)..... 21

Watson v. Fort Worth Bank & Trust, 487 U.S. 977 (1988) 15

Statutes:

28 U.S.C. § 1254(1)..... 1

28 U.S.C. § 1331 11

28 U.S.C. § 1343(a)(1) 11

28 U.S.C. § 1343(a)(3) 11

28 U.S.C. § 1343(a)(4) 11

28 U.S.C. § 1345 11

42 U.S.C. § 2000e-1, *et seq.* 11

42 U.S.C. § 2000e-2(a) 1

42 U.S.C. § 2000e-2(l) 23

Statutes:

Page

42 U.S.C. § 2000e-2(k) 1

42 U.S.C. § 2000e-2(k)(1)(A)(i) 1, 12

42 U.S.C. § 20003-2(k)(1)(A)(ii) 1

42 U.S.C. § 2000e-2(k)(1)(B)(i) 1

42 U.S.C. § 2000e-2(k)(1)(B)(ii) 1

42 U.S.C. § 2000e-6 ii

42 U.S.C. § 2000e-6(b) 11

29 C.F.R. § 1607, *et seq.* 1

29 C.F.R. § 1607.3(A) 23

29 C.F.R. § 1607.5(B) 15, 21

29 C.F.R. § 1607.5(C) 21

29 C.F. R. § 1607.5(H) 22

74 Pa. Con. Stat. Ann. § 1701, *et seq.* ii

Miscellaneous:

137 Cong. Rec. S15,276 (daily ed. Oct. 25, 1991) 12

Principles for the Validation and Use of Personnel Selection Procedures (1987) 22

Law Reviews and Treatises:

C. Alexander, *The Defeat of the Civil Rights Act of 1990: Wading Through the Rhetoric in Search of Compromise*, 44 Vand. L. Rev. 595 (1991) 13

R. Alioto, *Disparate Impact Discrimination Under the 1991 Civil Rights Act*, 45 Rutgers L. Rev. 1011 (1993) 12, 13

P. Bazer, *Mutable Characteristics and the Definition of Discrimination under Title VII*, 20 U.C. Davis L. Rev. 769 (1987) 13

TABLE OF AUTHORITIES—*(Continued)*

Law Reviews and Treatises: Page

K. Browne, *The Civil Rights Act of 1991: A Quota Bill, "A Codification of Griggs, a Partial Return to Wards Cove, or All of the Above*, 43 Case W. Res. L. Rev. 287 (1993) 13

L. Lye, *Comment; Title VII's Tangled Tale: The Erosion and Confusion of Disparate Impact and the Business Necessity Defense*, 19 Berkeley J. Employment & Lab. L. 315 (1998) 13

M. Rossein, *Disparate Impact After the Civil Rights Act of 1991: Restoring The Job Performance Standard*, (PLI, No. 1-14-5127-(1992) 16

A. Spiropolous, *Defining the Business Necessity Defense to the Disparate Impact Cause of Action: Finding the Golden Mean*, 74 N.C.L. Rev. 1479 (1996) 13

The Petitioner, Southeastern Pennsylvania Transportation Authority, respectfully prays that a writ of certiorari issue to review the judgment and opinion of the United States Court of Appeals for the Third Circuit, entered in the above entitled proceedings on June 29, 1999.

OPINIONS BELOW

The opinion of the Court of Appeals for the Third Circuit is reported at 181 F.3d 478 (3rd Cir. 1999), and is reprinted in Appendix A hereto, pp. 1a-45a, *infra.*

The memorandum decision of the United States District Court for the Eastern District of Pennsylvania (Clarence C. Newcomer, D.J.) has not been reported, however, it can be found at 1998 U.S. Dist. LEXIS 9388 (E.D. Pa. 1998). It is reprinted in Appendix B hereto, pp. 46a-162a, *infra.*

STATEMENT OF JURISDICTION

The jurisdiction of this Court to review the judgment of the Third Circuit, entered on June 29, 1999, is invoked pursuant to 28 U.S.C. §1254(1).

STATUTORY PROVISIONS INVOLVED

The Civil Rights Act of 1991, 42 U.S.C. §§2000e-2(a), 2000e-2(k), 2(k)(1)(A)(i)(ii), 2(k)(1)(B)(i)(ii) (reproduced herein at Appendix C); Uniform Guidelines for Employee Selection Procedures, 29 C.F.R. §1607, *et seq.*, (reproduced in part herein at Appendix D).

STATEMENT OF THE CASE

SEPTA is a regional mass transit authority operating in the five-county metropolitan area in and around Philadelphia. SEPTA operates a subway system, elevated rail lines, regional rail lines, bus lines and trolley lines. SEPTA's transit system transports more than 200,000 riders on a daily basis, and is the principle mass transit authority serving these areas.

The District Court found that in 1989, SEPTA's Transit Police Department was ineffective at controlling crime. Crime was rampant in SEPTA's transit system and its riders were being victimized by criminals. Significantly, at that time,

SEPTA's transit police department had no physical abilities standards or physical training program.

Faced with the fact that its patrol officers were not adequately protecting its riders, SEPTA implemented the following measures: (1) increased the number of its officers from 96 to 200; (2) introduced a "zone concept" of foot patrolling by decentralizing its officers into 8 zones; and (3) introduced physical abilities standards for its officers.

SEPTA hired Paul O. Davis, Ph.D. ("Davis"), a preeminent expert in the development of tests for protective service personnel, (i.e., police, fire fighters, and armed forces), to develop a validated aerobic capacity test. SEPTA advised Davis that its primary objective was to enhance the fitness, productivity and effectiveness of its police force and directed Davis to design a test targeting its goal to better protect its riders.

Davis performed a comprehensive job analysis of the SEPTA transit police officer position. In analyzing the position of a SEPTA transit police officer, Davis went on a twenty hour tour of SEPTA's transit system and conducted detailed information gathering sessions with SEPTA officers. Davis designed SEPTA's aerobic capacity test in light of the following information that he learned from SEPTA:

- 90% of SEPTA's transit police officers, unlike typical car-based police officers, patrol their beats *alone and on foot* for their entire eight hour shifts; the zone beats typically span 2-3 transit stations (5-8 city blocks);
- SEPTA officers patrol lengthy distances, encountering many stairs (30-50) and obstacles (i.e., turnstyles, barriers, etc.) with 14-26 lbs. of equipment;
- officers are routinely required to run in the performance of critical job tasks;

- officers are required to respond to officer assist calls, at a paced run, which are made when an officer is involved in, or about to become involved in, a potentially life-threatening situation;
- officers are required to respond to officer backup calls, which are made when an officer is confronted with a situation under control; however, the situation may escalate to an officer assist call, absent the presence of another officer; officers responding to backup calls run at a paced jog to the location;
- officers responding to officer backup or assist calls must arrive promptly at the location where assistance is needed, while maintaining enough reserve energy to engage in any arduous and physically demanding police duties that may be required;
- SEPTA's transit police department averages 4 running assist responses per zone per month; for the 8 zones, this totals approximately 32 running assists per month or *380 running assists involving potentially life-threatening situations per year*;
- for the 8 zones, SEPTA officers average approximately 160 running backups per month or approximately *1,920 running backups per year*;
- SEPTA officers often engage in running pursuits of criminals whom they must overtake, restrain and apprehend, often alone, before assistance can arrive.[1]

Davis determined the energy costs of the performance of physical tasks engaged in by SEPTA's officers to determine the aerobic capacity required to perform the critical tasks of the job.[2] Davis found that regularly performed tasks, such as pro-

---

1. The Third Circuit has erroneously stated that SEPTA officers run only *twice* per year. The trial record contains ample evidence of the total number of times that officers must run in performing their duties. Taking into account backups and assists, SEPTA's transit police department runs 2,300 times per year. This figure does not even take into account the number of times officers may have to run to pursue fleeing criminals. Thus, each officer averages at least *12* running events per year, not 2, as stated by the majority in footnote 23 (App. A, at p. 25a).

2. Aerobic capacity is the ability of the body to utilize oxygen during

viding backup or assistance to another officer at the next station, require 43.5 mL/kg/min.[3] Frequently, SEPTA officers, upon arrival at the scene of an emergency, must surmount 30-50 transit station steps. This job requirement demands a higher energy reserve from the officer; an officer would need to have an aerobic capacity of 54 to be able to promptly arrive at the scene of a call and ascend the typical flight of stairs in SEPTA'S transit system.

Davis' calculations showed that the aerobic capacity required for a transit officer to successfully perform many critical tasks of the position ranges from 42.5 to 54 depending on the circumstances of the running activity. Davis concluded that a SEPTA officer must have an aerobic capacity of 50 to successfully perform many commonly required critical tasks of the position.

Based on normative data on aerobic capacity levels for males and females, Davis believed that a standard of 50 could present a substantial obstacle for female applicants for the position of SEPTA transit police officer. Davis knew that women *could readily* attain a standard of 42.5. Consequently, to reduce the likelihood that his test would exclude a higher number of female applicants, Davis designated SEPTA's cut-off score to be 42.5.[4]

Based on his targeted cut-off score of 42.5, SEPTA implemented a physical abilities screening test that required successful applicants to run 1.5 miles in 12 minutes or less, which

---

sustained physical activities such as running, swimming and cycling.

3. Aerobic capacity is commonly measured in units of milliliters of oxygen per kilogram of body weight per minute ("mL/kg/min"). Hereinafter, the aerobic capacity levels will be discussed by number only and without the mL/kg/min measure.

4. The Third Circuit, in footnote 19 (App. A, at p. 23a), criticized Davis for rejecting officer-estimates that officers need to be able to run one mile in 11.78 minutes. Davis rejected this pace because it translates to an aerobic capacity of 33.5, which is less than the aerobic capacity of a sedentary female (37). A level of 33.5 is not appropriate for a safety-sensitive job that entails critical, life-and-death tasks, dependent on officer aerobic capacity, ranging from 42.5 to 54. At trial, plaintiffs' expert in test development, Dr. Sheldon Zedeck, conceded that the test developer's judgment can and should be exercised by the person validating the test.

translates to an aerobic capacity of 42.5. The District Court found that SEPTA's running test does, in fact, measure whether an applicant has the level of aerobic capacity that Davis identified as necessary to perform many critical tasks for the patrol officer position and to be able to successfully perform his or her duties.

SEPTA also introduced an incumbent testing program to increase the aerobic capacity of its incumbent patrol officers to the standards expected of its applicants within a reasonable time period. SEPTA originally disciplined incumbent officers who did not succeed in passing an individualized program of gradual improvement of their aerobic capacities. This program was undercut when the disciplined incumbents filed a labor arbitration grievance complaining about the discipline and prevailed. The arbitrator ruled that disciplining incumbents for failing physical abilities goals violated the collective bargaining agreement. Despite SEPTA's inability to discipline or fire its incumbents, 86% of the incumbents hired *prior* to SEPTA's implementation of its aerobic capacity test attained the 42.5 standard.

Because 86% of the incumbent officers have attained an aerobic capacity of at least 42.5, and because 97% of the incumbent officers hired after 1991 still meet or exceed the 42.5 standard, SEPTA has been able to further its goal of better protecting its riders. Notably, since the inception of the aerobic capacity test in 1991, SEPTA has reduced crime in its transit system by 70%.

Five female applicants that failed SEPTA's aerobic capacity test in 1993 commenced this action. Initially, these plaintiffs filed their charge of discrimination with the Pennsylvania Human Relations Commission ("P.H.R.C.") and the Equal Employment Opportunity Commission ("E.E.O.C.") in their individual capacities and as class representatives. On December 11, 1996, the E.E.O.C. issued a Right to Sue notice after the P.H.R.C. made a finding of probable cause based on the adverse impact of SEPTA's aerobic capacity test on female applicants.

In February of 1997, the Attorney General of the United States commenced a separate action against SEPTA, claiming

that SEPTA was engaging in a pattern and practice of discrimination by using a test that was not predictive of successful job performance. Subsequently, these suits were consolidated for discovery and trial.

At trial, female applicants who passed SEPTA's aerobic capacity test emphasized that SEPTA transit officers must be physically fit to meet the rigorous demands of the job.[5] SEPTA officials corroborated this point. In support of SEPTA's use of the aerobic capacity test, SEPTA introduced testimony from the following expert witnesses: Davis, who devised the test; Lynda Ransdell, Ph.D., an expert physiologist with a speciality in women's fitness; Robert Moffat, Ph.D., a physiologist (who performed post-validation studies); Bernard Siskin, Ph.D., a statistician (who performed post-validation studies); and Norman Henderson, Ph.D., an industrial and organizational psychologist.

Dr. Moffat conducted validation studies with data derived from real-to-life officer assist and officer backup simulations performed by SEPTA officers. Based on Moffat's studies, the District Court found that officers with aerobic capacities of less than 45 are compromised in performing arduous tasks (arrests, apprehensions and combative struggles) after a 3-5 block paced run that is routinely required of SEPTA officers. Based on his studies and data, Moffat opined that SEPTA's standard of 42.5 was very conservative. The District Court found that Moffat's study proved that officers with aerobic capacities of less than 42.5: (1) may not be able to arrive on time to help the fellow officer in need of assistance or back up;

---

5. The requirements of SEPTA's aerobic capacity test were made known to applicants well in advance of the administration of the test. The successful female applicants all prepared for the test by training in accordance with the instructions contained in SEPTA's test information packet. For example, Officer Bernadette Rodier, a 39 year-old woman who worked as a waitress for 15 years, started running a few times a week to prepare for the test (App. B, at p. 84a). In contrast, the named plaintiffs generally failed to train for the test and often walked during the administration of the aerobic capacity test. (App. B, at pp. 79a-83a); *see also*, videotaped portions of plaintiffs' aerobic capacity test: Heats 1, 3, 8 and 18— Defendant's trial exhibit 108, one copy of the tape has been filed. *See* Letter filing tape, at App. F.

(2) may not be able to arrive in time to apprehend the perpetrator; or (3) once the officer arrived, the officer was likely to be without adequate energy reserves to be effective in the arduous activities that may be required of the officer. The District Court found that Moffat's studies proved that:

- **Officers with aerobic capacities of 45 or more had only a nominal decrement (9-10%) in their ability to apprehend, arrest, and engage in struggles after jogging or running to an officer backup or assist**
- **Officers with aerobic capacities of less than 45 suffered severe decrements (30%) in their ability to apprehend, arrest, and engage in struggles.**

SEPTA also introduced the continuing validation studies of Bernard Siskin, a renowned statistician. Siskin's studies conclusively proved, and the District Court found, that SEPTA's aerobic capacity requirement of 42.5 significantly correlates with and predicts *more* arrests of felony offenders (Part I Crimes), *more* overall arrests in all categories of crimes, and *higher* arrest rates.[6] The arrest studies proved the following:

- **Officers always at 42.5 made 3 times (151%) the number of Part I arrests than officers always below 42.5**
- **Officers at or above 42.5 have a 32% arrest rate advantage for Part I crimes**
- **Officers at or above 42.5 have a 14% overall arrest rate advantage**
- **A regression analysis proved that from 1991-96, SEPTA could have achieved 469 additional arrests, 70 of which were Part I arrests (felony arrests), if the aerobic capacity of all of its officers had been 42.5 or more during this period; these additional arrests reflect a 10% increase in Part I arrests and a 4% increase in overall arrests**
- **SEPTA would gain a ½% increase in Part I arrests for every increase of 1 ml of aerobic capacity; this relationship was linear, i.e., the higher the aerobic capacity of**

6. Dr. Henderson testified that reputable studies in the area of law enforcement have identified arrests, felony arrests and commendations as important and appropriate criterions for the measurement of patrol officer performance.

the officer, the higher the officer's arrest rate, number of felony arrests and overall arrests.[7]

Siskin also studied 207 commendations awarded to officers for valor, bravery, heroism and meritorious actions from 1994-96. The study demonstrated, and the District Court found, the following facts:

- **96% of the commendations were awarded to officers with an aerobic capacity greater than or equal to 42.5 (the commended officers averaged 46.6)**
- **Officers at 42.5 or more are 10 times as likely to be awarded commendations than officers at 39 or less**
- **Of the 207 commendations studied, 198 involved an arrest and 116 had an explicit reference in the commendation to a foot pursuit or other physical exertion**
- **The mean aerobic capacity for the commended officers (47), when compared with the entire officer population (44), was higher at a statistically significant level (.01).**

The statistical evidence from Siskin's studies showed, and the District Court found, that perpetrators commiting crimes in SEPTA's transit system maintain aerobic capacities at or above 48. In reaching this conclusion, Siskin studied 953 per-

---

7. The Third Circuit, in footnote 23, (App. A, p. 25a), attempts to debunk and diminish linearity, a well accepted principle of test validation. In order to sidestep this critical and well settled principle of test validation, the Circuit states, without support, that for linearity to have relevance, the test must encompass the "totality of a job's duties." The majority opinion also demonstrates a fundamental lack of understanding of statistics. Pointing to a *single* data point on a trial exhibit, the majority claims that the relationship between aerobic capacity and arrests is not strictly linear. The Circuit ignores the fact that data will rarely, *if ever*, be "strictly" linear due to random chance variation. Linearity is a *demonstrated trend* and, here, the arrest data shows a statistically *significant* linear relationship with aerobic capacity. Similarly, the majority's comparison of officers with aerobic capacities of less than 37 and their overall arrests, to the overall arrests made by officers with aerobic capacities of more than 48, is an inappropriate comparison which ignores the empirical evidence of the *733 remaining data points*, or 64% of the data in that Table. If the Circuit examined the relationship between officers with aerobic capacities of less than 37 and their *Part I arrests*, and officers with aerobic capacities of more than 48 to their *Part I arrests*, it would have found that the 48 category had *twice* the number of Part I arrests (with statistical significance at less than .01).

petrators arrested for Part I crimes over a six-year period to determine their aerobic capacities, and analyzed the data in light of the perpetrators' sex, age and race. Siskin's studies showed, and the District Court found that:

- **The average age of perpetrators is 26; and their average aerobic capacity is 48**
- **Nearly 52% of the perpetrators have an aerobic capacity of 48 or more**
- **Only 27% of the perpetrators have an aerobic capacity of less than 42.5.**

From a study of 382 Part I arrests from 1994-96, SEPTA proved, and the District Court found, that the arresting officers had a mean aerobic capacity of *46.8*, while the criminals arrested for Part I crimes averaged *48*.[8] Siskin's studies showed, and the District Court found that:

- **94% of the officers arresting Part I perpetrators had an aerobic capacity greater than or equal to 42.5 (the average for this category of officers was 46.8)**
- **Officers apprehending Part I criminals had aerobic capacities higher than that of the SEPTA patrol officer population at, a statistically significant level (.01).**[9]

---

8. Incredulously, in footnote 21, (App. A, at pp. 22a-24a), the Third Circuit complains that serious crimes (Part I crimes) account for *only* 7% of all arrests and, therefore, represent only "a small aspect of the job." This completely misses the point. In focusing on this percentage, the majority erroneously and intentionally equates "frequency" with "importance." It cannot be seriously argued that an officer's ability to make Part I arrests when necessary is *not* a critical task.

9. The Third Circuit, in footnote 18 (App. A, at p. 21a), inappropriately criticizes the perpetrator study, claiming that it is unlikely that perpetrators of felony crimes will engage armed officers in "an aerobic contest." This ignores the fact that perpetrators of felonies may resist apprehension or flee from arrest. The officer's last resort is to use deadly force. Contrary to the Circuit's speculation, all of Siskin's studies demonstrate that a clear and significant correlation exists between an officer's aerobic capacity and his or her ability to arrest felons.

Dr. Ransdell, a physiologist specializing in women's fitness, testified that with minimal training women would be able to pass SEPTA's aerobic capacity test. Ransdell testified that:

- **50% of all women starting at a sedentary aerobic capacity of 35 to 37 can train to pass SEPTA's 1.5 mile aerobic capacity test in 8 weeks**
- **10% of all women between 20-29 years of age can meet SEPTA's standards without *any* training**
- **40% of all women between 30-39 years of age can meet SEPTA's standard with only moderate training**
- **Female runners *over the age of 50* generally have aerobic capacities from 44 to 50.**

At trial, the plaintiffs failed to introduce any statistical studies. Rather, the plaintiffs offered the testimony of: Dr. Sheldon Zedeck, an expert on test development; representatives from dissimilar police forces; and Dr. William McArdle, a physiologist who has never designed or validated any physical abilities tests. At trial, as a "less restrictive alternative," McArdle proposed that the aerobic capacity test be set at 42.5 for males and 36 for females. Significantly, Zedeck, plaintiffs' own expert on test development disavowed this gender-based proposal. Zedeck testified that he would never design a test with differing standards for males and females for the same job. The plaintiffs called witnesses from other police jurisdictions to advance the proposition that because other departments had no physical abilities tests, physical abilities testing for a police officer position must be unnecessary. Notably, the other departments that testified never undertook a job study of their own force and never implemented a validated physical abilities test. Ironically, their testimony ultimately supported SEPTA's position; the other police departments admitted that physical fitness is important to the job and that arrests, a criterion studied by SEPTA, is a critical aspect of a patrol officer's duties.

BASIS FOR FEDERAL JURISDICTION AND DECISIONS BELOW

Invoking federal jurisdiction for Civil Action No. 97-0593 (the named plaintiffs/class representatives) pursuant to 42 U.S.C. §2000e-1 *et seq.*, and 28 U.S.C. §§1331, 1343 (a)(1), (a)(3) and (a)(4) and for Civil Action No. 97-1161 (the United States) pursuant to 42 U.S.C. §2000e-6(b) and 28 U.S.C. §§1343(a)(3) and 1345, the Respondents brought this suit in the Eastern District of Pennsylvania. On June 25, 1998, the District Court rendered a verdict in favor of SEPTA, the Petitioner. *See* App. B, at p. 46a, *infra*.

The District Court, after a 2½-week bench trial, issued a 162-page opinion delineating the scientific, professional and statistical evidence that showed that transit officers must have an aerobic capacity of at least 42.5 to meet the demands of the transit patrol officer position. The District Court, concurring with SEPTA's experts, found that aerobic capacity correlates with and predicts important criterions of field performance, i.e., more arrests, higher arrest rates, and the award of commendations. The District Court also found that this relationship was linear, i.e., increasing levels of aerobic capacity translated into better and more effective patrol officer performance (more arrests, higher arrest rates), all to the benefit of public safety.

In accordance with the Civil Rights Act of 1991, the District Court applied the criteria of 'job-relatedness' and 'business necessity' found in this Court's cases prior to *Wards Cove Packing Co. v. Atonio*, 490 U.S. 62 (1989). The District Court found that SEPTA's applicant aerobic capacity test was 'job-related' and met the standard of 'business necessity.' Additionally, the District Court rejected the plaintiffs' proposed alternatives because they did not equally serve SEPTA's demonstrated business requirements.

On June 29, 1999, a divided panel of the Third Circuit entered a judgment and opinion vacating the District Court's judgment and remanding the case back to the Eastern District for an evaluation of the evidence under a new standard of 'business necessity.' The Third Circuit did not challenge the District Court's finding that SEPTA met its burden of proving

that its aerobic capacity test was 'job-related.' Articulating a new standard of 'business necessity,' the Third Circuit held that, pursuant to the Civil Rights Act of 1991, when plaintiffs show that a test has disparate impact, the employer must demonstrate that it selected a cut-off score for an entry level employment examination that measures only the *"minimum qualifications"* necessary to perform the job. SEPTA did not file a Petition for Rehearing.

REASONS FOR GRANTING THIS WRIT

I. SINCE THE SUPREME COURT HAS NOT EXAMINED THE 'BUSINESS NECESSITY' TEST SINCE *WARDS COVE*, OR AFTER THE ENACTMENT OF THE CIVIL RIGHTS ACT OF 1991, THIS COURT SHOULD NOW CLARIFY THE 'BUSINESS NECESSITY' STANDARD AND SETTLE THE CONFLICT IN THE CIRCUITS

Congress' presumption that the term 'business necessity' in the Civil Rights Act of 1991 connotates a uniform meaning[10] is refuted by this Court's various formulations of 'business necessity' and by the various inconsistent standards of 'business necessity' articulated by several Circuit Courts. This lack of uniformity both invites and demands clarification from

---

10. Section 105(a) requires the employer to demonstrate that the challenged practice is "job-related for the position in question and consistent with business necessity." *See* 42 U.S.C. §2000e-2(k)(1)(A)(i). The statute itself provides no definition of 'business necessity.' Instead, the approved legislative history of the Act states that "business necessity" must reflect this Court's pronouncements in *Griggs* and other pre-*Wards Cove* cases. *See* 137 Cong. Rec. S15,276 (daily ed. Oct. 25, 1991). This statutory directive provided no guidance to the Circuits because it directed them back to standards that were already unclear. This is evidenced by the fact that some Circuits equated the standards of 'job relatedness' and 'business necessity' and deemed them to be interchangeable. Additionally, the confusion by and between the Circuits was reflected by other Circuits articulating an "absolute necessity requirement for a job qualification" and by other Circuits applying a sliding scale approach based on the complexity and demands of the job. *See* R. Alioto, *Disparate Impact Discrimination Under the 1991 Civil Rights Act*, 45 Rutgers L. Rev. 1011, 1029 (1993) (collecting cases).

this Court.[11] Significantly, the following invitation has recently been extended to this Court by Judge Weis — who dissented from, and disagreed with, the Third Circuit's radical departure from all Title VII disparate impact cases, including the decisions of this Court and all other Circuit Courts:

> The Supreme Court's articulations of the appropriate standards are far from clear. Phrases such as "business necessity," "demonstrable relationship to successful performance of the job," "manifest relationship to the employment in question," "genuine business needs," and "essential to good job performance," have been used interchangeably. These varying formulations bring to mind Justice Holmes' observation, "A word is not a crystal, transparent and unchanged, it is the skin of a living thought and may vary greatly in color and content according to the circumstances and the time in which it is used."

---

11. The absence of a definition of 'business necessity' in the Civil Rights Act has caused courts and scholars alike to attempt to decipher (unfortunately by resorting to supposition) what the pre-*Wards Cove* standard was and what the present standard is. *See, e.g.*, R. Alioto, *supra* (noting that Circuits, pre-*Wards Cove* had required either 'job-relatedness' or 'business necessity' and that cases which require proof of 'job-relatedness' and a reasonable need for the challenged practice would be in accord with the 1991 Act and Supreme Court precedent); A. Spiropolous, *Defining the Business Necessity Defense to the Disparate Impact Cause of Action: Finding the Golden Mean*, 74 N.C.L. Rev. 1479 (1996) (analyzing pre-*Wards Cove* Supreme Court cases based on differences in job characteristics, including safety concerns); L. Lye, *Comment: Title VII's Tangled Tale: The Erosion and Confusion of Disparate Impact and the Business Necessity Defense*, 19 Berkley J. Employment & Lab. L. 315 (1998) (challenged practice must be a reasonable predictor of effective performance of job duties, defined in light of important business goals); K. Browne, *The Civil Rights Act of 1991: A Quota Bill, A Codification of Griggs, A Partial Return to Wards Cove, or All of the Above*, 43 Case W. Res. L. Rev. 287 (1993) ('business necessity' has the same meaning as *Wards Cove* phrase "serves in significant way"); P. Bazer, *Mutable Characteristics and the Definition of Discrimination Under Title VII*, 20 U.C. Davis L. Rev. 769 (1987) (both the Supreme Court and lower court rulings offer a confusing patch work of seemingly conflicting standards); C. Alexander, *The Defeat of the Civil Rights Act of 1990: Wading Through the Rhetoric in Search of Compromise*, 44 Vand. L. Rev. 595 (1991) (discussing the rejected bills and quota issue).

> *Towne v. Eisner*, 245 U.S. 418, 425 (1918). ***Until the Supreme Court re-examines the subject, however, courts will continue to struggle with the often inconsistent phraseology employed in Griggs and its progeny.*** The definition and application of the appropriate standard for business justification will depend on the context in which it is raised . . . (App. A, at p. 35a) (Emphasis added).

Unless and until this Court defines the standard of 'business necessity,' courts will continue to craft and apply varying, ambiguous and inconsistent definitions of 'business necessity.' The varying and inconsistent standards of 'business necessity' among the Circuits highlight that the Civil Rights Act's goal of yielding uniform results will continue to be unfulfilled, unless this Court intercedes. Fortunately, this case presents this Court with the opportunity to resolve this controlling question of law. Clarification from this Court as to what constitutes 'business necessity' is both wanted and required.

The outcome of an action should not be contingent on where it was filed. Likewise, employers should not compelled to design and defend their applicant screening tests without knowing the type or quantum of evidence required to meet this ambiguous standard. The Third Circuit's contribution of a new standard of 'business necessity' adds yet another formulae to the ever-growing delineations of 'business necessity.' Clearly, the construction of a legal test should not, and need not, be left to supposition.

II. THE "MINIMUM QUALIFICATIONS" TEST OF 'BUSINESS NECESSITY' CONFLICTS WITH EVERY SUPREME COURT AND CIRCUIT COURT FORMULATION OF 'BUSINESS NECESSITY' UNDER TITLE VII

The Third Circuit's enunciation of a standard that is a radical departure from this Court's decisions defining 'business necessity,' and from *every* Circuit Court's articulation of the test of 'business necessity' under Title VII, provides this Court with an independently compelling reason to review this case.

Contrary to the Third Circuit's claim that its novel "minimum qualifications" test of 'business necessity' is required by a faithful reading of *Griggs*, its standard is unsupportable, and in fact, is contradicted by each pre-*Wards Cove* case decided by this Court, *including Griggs*. *See Griggs v. Duke Power Co.*, 401 U.S. 424, 436 (1971) (defining 'business necessity' as a selection criterion that is *"related to job performance"* and *"demonstrably a reasonable measure of job performance"*) (Emphasis added); *New York City Transit Authority v. Beazer*, 440 U.S. 568, 587 (1979) and *Dothard v. Rawlinson*, 433 U.S. 321, 329 (1977), (*citing Griggs, supra*, 401 U.S. at 432) (holding that the 'business necessity' standard is satisfied when the selection criteria is shown to bear a *"manifest relationship to the employment in question"*) (Emphasis added); *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 431 (1975) (finding that the 'business necessity' standard is met where the employer shows that the test is *"predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job . . . for which candidates are being evaluated"*) (Emphasis added); *Watson v. Fort Worth Bank & Trust*, 487 U.S. 977 (1988) (plurality decision), (holding that 'business necessity' requires that employment tests *directly relate to "the ability to perform the job effectively"*) (Emphasis added). *See also* 29 C.F.R. §1607.5(B) (a test should consist of empirical data demonstrating that the selection procedure is predicative of, or significantly correlated with, important elements of job performance).

Although this Court has articulated varied formulations of 'business necessity,' none of its decisions authorize or support the Third Circuit's unique requirement that a testing cut-off score must only measure the "minimum qualifications" for an entry level position. The Third Circuit's claim that its new "minimum qualifications" standard is implicit in, or required by *Griggs* and its progeny, is disingenuous at best, and is contradicted by the very precedent it relies upon.

Similarly, the Third Circuit's unprecedented standard is inconsistent with every other Circuit Court decision rendered

prior to *Wards Cove*,[12] and after the enactment of the Civil Rights Act of 1991.[13] The Third Circuit's contribution of a *new*

---

12. As of 1992, there were more than 275 cases decided prior to *Wards* Cove which discussed the standard of 'business necessity.' *See* M. Rossein, *Disparate Impact After the Civil Rights Act of 1991: Restoring the Job Performance Standards*, at Appendix A, PLI # H4-5127-1992. While it is not practical to list all of these cases herein, an examination of these cases reveals that *no* Circuit has gone as far as the Third Circuit in articulating as onerous of a burden on the employer to prove 'business necessity.' For example, in *Robinson v. Lorillard Corporation*, 444 F.2d 791, 798 (4th Cir. 1971), the Fourth Circuit, interpreting *Griggs*, held that "the [employment test] must have a manifest relationship to the employment in question." In *Pettway v. American Cast Iron Pipe Co.*, 494 F.2d 211, 244-45 (5th Cir. 1974), the Fifth Circuit, interpreting *Griggs*, held that the employer must demonstrate that its selection device has a "manifest relationship to the employment in question" and that the "employer must demonstrate that the selection device is 'valid.' " In *Harless v. Duck*, 619 F.2d 611, 616 (6th Cir. 1980), the Sixth Circuit, interpreting *Dothard*, construed the "business necessity" defense as requiring the employer to demonstrate that "the test bears a manifest relationship to successful and efficient job performance." In *Contreras v. City of Los Angeles*, 656 F. 2d 1267 (9th Cir. 1981), the Ninth Circuit, interpreting *Griggs*, *Dothard*, and *Albermarle*, held that an employer meets the 'business necessity' standard where the employer demonstrates a significant relationship between the challenged selection device or criteria, and important elements of the job, and there need not be a perfect fit.

13. Other Circuits interpreting 'business necessity' after the enactment of the Civil Rights Act of 1991, do not even remotely suggest that the defense requires proof that the employment test measures the lowest level of qualifications to successfully perform the job. For instance, in *NAACP v. Town of East Haven*, 70 F.3d 219, 225 (2d Cir. 1995)(*citing Albemarle*, 422 U.S. at 425 and *quoting Griggs*, 401 U.S. at 432), the Second Circuit held that once the burden shifts to the employer, the employer must come forward with evidence to show that the test has "a manifest relationship to the employment in question." In *Smith v. City of Des Moines*, 99 F.3d 1466, 1471 (8th Cir. 1996), the Eighth Circuit held that the City needed to establish that: (1) the requirement has a "manifest relationship to the employment in question" (*citing Griggs*, 401 U.S. at 432) and (2) "it is necessary to safe and efficient job performance" (*citing Dothard*, 433 U.S. 321, at footnote 14). In *Allen, et al v. Energy Corporation, et al.*, 181 F.3d 902 (8th Cir. 1999), the Eighth Circuit affirmed the dismissal of a race discrimination action, emphasizing that "Title VII forbids the use of employment tests that are discriminatory in effect unless the employer meets the burden of showing that any given requirement has a manifest relationship to the employment in question. . . . The standard of proof articulated in the *Griggs* case and codified in 42 U.S.C. §2000e - 2(k) states that where the plaintiff has made a *prima facie*

and more onerous definition of 'business necessity' moves the Circuits even further away from the goal of uniformity, and highlights the immediate need for long awaited intervention from this Court.

III. THE THIRD CIRCUIT'S REFUSAL TO RECOGNIZE PUBLIC SAFETY AS A CONSIDERATION IN ASSESSING THE 'BUSINESS NECESSITY' OF A TEST CONFLICTS WITH DECISIONS OF THIS COURT, AND WITH SIX OTHER CIRCUIT COURTS AND ADVERSELY IMPACTS THE SAFETY OF MILLIONS OF CITIZENS

The Third Circuit's new and unsupportable "minimum qualifications" standard requires employers to *lower* occupational standards for entry level applicants to the *lowest common denominator,* even in a safety-sensitive position and usurps an employer's right to set the qualifications for its employees. As Judge Weis pointedly observed in his dissenting opinion:

> [r]educing standards towards the lowest common denominator is particularly inappropriate for a police force. Undoubtedly, candidates who fail the running test — female or male — may have other qualities of particular value to SEPTA, but they must possess the requisite aerobic capacity as well. No matter how laudable it is to reduce job discrimination, to achieve this goal by lowering important public safety standards presents an unacceptable risk. *See* App. A, at p. 41a.

---

case of disparate impact from a selection procedure the burden then shifts to the employer to justify the procedure by showing that it is related to safe and efficient job performance." In *Association of Mexican-American Educators,* 181 F.3d 902 (9th Cir. 1999), the Ninth Circuit held that "the employer must demonstrate by professionally acceptable methods that the selection device is *predictive of or significantly correlated with* the element of work behavior identified . . .". In *Fitzpatrick v. City of Atlanta,* 2 F.3d 1112, 1117 (11th Cir. 1993), the Eleventh Circuit, interpreting the pre-*Wards Cove* cases, i.e. *Beazer* and *Dothard,* stated that the employer has the burden of showing that the challenged action is demonstrably necessary to meeting a goal of a sort that, as a matter of law, qualifies as an important business goal for Title VII purposes.

The Third Circuit's "minimum qualifications" standard of 'business necessity' conflicts with this Court's precedent[14] and six Circuits that consider public safety when assessing whether an employer screening to fill safety-sensitive positions has satisfied the standard of 'business necessity.' The Third Circuit's decision uniquely imposes an even higher and more exacting burden of proof on the employer, and deems the consideration of public safety to be irrelevant to its analysis.[15] Six other Circuits have held that safety considerations must be considered when evaluating the validity of applicant testing in safety-related jobs — precedent that imposes a lighter burden on an employer to meet 'business necessity' where public safety is at issue. *See Davis v. City of Dallas*, 777 F.2d 205 (5th Cir. 1985) (in context of applicant testing for position of police officer, emphasizing that the degree of public risk and responsibility which the job entails warrants imposing a lighter burden to prove 'business necessity'); *Chrisner v. Complete Auto Transit, Inc.*, 645 F.2d 1251, 1262-63 (6th Cir. 1981) (requiring lighter burden to meet 'business necessity' test in disparate impact case, due to public safety concerns, in the context of applicant screening requirement for long-distance truck drivers); *Hodg-*

---

14. *See Beazer, supra*, 440 U.S. 568 (1979) (emphasizing that public safety must be considered where the position is attended by unusual hazards and must be performed by persons of maximum alertness and competence); *Washington v. Davis*, 426 U.S. 229 (1976) (lighter burden required in public safety context in evaluating fit between test and job performance). Six Circuits have relied on *Beazer* and *Davis*, as well as other Supreme Court cases, to support the propositions that where public safety is at issue a lighter burden is imposed on the employer in demonstrating 'business necessity' and that safety considerations are to be taken into account.

15. In contrast, in footnote 16 of its majority opinion, (App. A, pp. 19a-20a), the Third Circuit flatly rejects, as inconsistent with the Civil Rights Act of 1991, consideration of public safety when assessing the validity of a test for a safety-sensitive position. In his strong and incisive dissent, Judge Weis not only disagrees, but believes that where public safety is at issue, "[d]ue deference should be afforded to the experience of specialized employers in setting appropriate requirements for safety-sensitive positions." *See Id.* at p. 44a. Judge Weis bases this conclusion on his reasoned analysis of this Court's pre-*Wards Cove* cases addressing safety, i.e., *Beazer, supra.*, and Circuit Court pre-*Wards Cove* cases, that specifically endorse a lighter burden for positions relating to public safety. *See Id.* at pp. 35a-38a.

son v. Greyhound Lines, Inc., 499 F.2d 859, 862-63 (7th Cir. 1974) (finding age to be a proxy for physical fitness and imposing lighter burden to find bona fide occupational qualification in the context of long distance bus drivers, where safety concerns are present); *Smith v. City of DeMoines*, 99 F.3d 1466, 1473-74 (8th Cir. 1996) (upholding use of aerobic capacity requirement in fire fighting context; holding that where public safety and the safety of fellow fire fighters are at stake, the lack of a precise fit between the test and the job requirement is not fatal to a defense of 'business necessity'); *Harris v. Pan American World Airways, Inc.*, 649 F.2d 670 (9th Cir. 1980) (in disparate impact case, recognizing the importance of public safety, in context of flight attendant position with commercial airline); *Spurlock v. United Airlines, Inc.*, 475 F.2d 216, 219 (10th Cir. 1972) (requiring lighter burden to meet 'business necessity' test in disparate impact case because of public safety concerns in the context of applicant testing of commercial airline pilots); *Fitzpatrick v. City of Atlanta*, 2 F.3d 1112, 1119-20 (11th Cir. 1993) (in disparate impact case, recognizing the importance of public safety and safety of fellow fire fighters).

The Third Circuit's articulation of a standard of 'business necessity' that intentionally ignores public safety considerations, despite the obvious — that the most critical job tasks of a SEPTA transit police officer are to protect the public and to protect fellow officers is, at best, misguided. This Court should exercise its discretion to review the Third Circuit's standard of 'business necessity' because it is expressly in conflict with decisions by the Fifth, Sixth, Eighth, Ninth, Tenth and Eleventh Circuits. Further, a declaration from this Court is required to address the split among the Circuits as to whether the burden of proof in a disparate impact case, in view of the Civil Rights Act of 1991, is the same for safety-sensitive positions as it is for jobs that do not entail critical tasks affecting public safety.

IV. THE THIRD CIRCUIT'S "MINIMUM QUALIFICATIONS" STANDARD PREVENTS EMPLOYERS FROM UPGRADING THE PERFORMANCE OF THEIR WORKFORCE THROUGH APPLICANT TESTING AND SETS ENTRY LEVEL JOB PERFORMANCE AT THE LOWEST COMMON DENOMINATOR FOR THOUSANDS OF EMPLOYERS

No decision of this Court authorizes the Third Circuit to articulate a standard of 'business necessity' that requires an employer to settle for mediocre job performance, or that prohibits an employer from upgrading entry level job performance through applicant testing. The Third Circuit's tortured reading of this Court's precedent to reach its outcome driven result constrains employers to the performance level of the least effective applicant and least effective incumbent employee.[16] When stripped of its rhetoric, the Third Circuit's 'minimum qualifications only' standard prohibits the use of validated pre-employment tests to raise productivity and upgrade job performance. This outcome is particularly egregious where the position impacts on public safety.[17] Contrary

---

16. In *Albermarle Paper Co. v. Moody*, 422 U.S. 405 (1975), this Court held that tests that result in disparate impact are impermissible unless shown, by professionally acceptable methods, to be predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated. Where SEPTA has proven, and the District Court has found, that increasing levels of aerobic capacity will result in corresponding increases in the number of overall arrests, felony arrests and arrest rates (i.e., linearity is shown), this standard is clearly satisfied.

17. The Third Circuit suggests that SEPTA's aerobic capacity test violates the Civil Rights Act of 1991 and the Uniform Guidelines on Employee Selection, because it does not test for "qualities that fairly represent the totality of a job's responsibilities." This notion is contrary to case law, the Uniform Guidelines and intentionally ignores that SEPTA's aerobic capacity test measures whether applicants will be able to perform the most critical tasks of the job – protecting the public and assisting fellow officers. If SEPTA's officers cannot perform these critical safety-sensitive tasks, it is largely irrelevant that he or she may be able to meet some or all of the other demands of the job. Significantly, the Third Circuit does not, and cannot, point to any legal precedent that would require SEPTA to test for the totality of the job's responsibilities. *See Guardians Association of New York City*

to the Third Circuit's 'opinion,' there can be no "minimum qualifications" where public safety is at issue and where increasing levels of aerobic capacity are shown to result in increases in the number of felony arrests, overall arrests and arrest rates. The concept of "minimum qualifications" is nonsensical in the face of a linear relationship between SEPTA's test's predictor—aerobic capacity—and these important criterions of patrol officer performance.[18]

Police Department v. Civil Service Commissions, 630 F.2d 79 (2d Cir. 1980) (test need *not measure all aspects, regardless of significance, in their exact proportions); Police Officers v. City of Columbus*, 916 F.2d 1092, 1109-10 (6th Cir. 1990) (*citing Guardians Association of New York City Police Department, supra*, flatly rejecting argument that lieutenant exam violates Title VII because it does not measure attributes in proportion to their importance and frequency of use in the performance of the job)); *Cleghorn v. Herrington*, 813 F.2d 992 (9th Cir. 1987) (finding that to better protect the security of nuclear power plants, *an employer is entitled to implement a physical fitness test that measures whether security guards will be able to effectively respond to emergency situations, even though they do not routinely occur, because if emergency situations do occur, effective responses are crucial to security).* The Third Circuit's intentional dismissal of the linear relationship between an officer's aerobic capacity and the number of arrests an officer can make, and an officer's capacity to effectively respond to officer calls for backup and assistance — the most critical job tasks — is disturbing in light of *Cleghorn*'s recognition that testing for critical safety tasks, regardless of their frequency, is appropriate. *See* 29 C.F.R. §§1607.5(B) and (C). *See also Beazer, supra,* and *Washington v. Davis, supra,* where this Court expressly approved of defining job requirements in light of important safety considerations. Further, it is incomprehensible that the Third Circuit believes that a 'more is better' concept is acceptable in a typing test context, but not in the context of safeguarding the public and fellow officers. (*See* App. A, at pp. 22a, fn. 19, and 25a, fn. 23).

18. The Third Circuit's 'minimum requirements' standard can only be understood as the product of the Third Circuit's dismissal of acknowledged principles of testing. To reach the "minimum qualifications" test of 'business necessity,' the Third Circuit substituted its judgment for that of the testing experts and expert literature concerning the import of a linear relationship and test utility. Once linearity is demonstrated, an employer may select where to set its cut-off score, because each additional increment of the predictor (aerobic capacity) yields better performance (more felony arrests, overall arrests and higher arrest rates). Even the plaintiffs' expert on selection devises, Dr. Zedeck, acknowledged that with a valid linear test: (1) the higher the cut-off score, the higher the utility of the test; (2) the cut-off score may be set at a gradation of performance above that which is merely satisfactory; and (3) selecting a cut-off score that excludes more women than

Clearly, the Civil Rights Act of 1991 did not eradicate an employer's right to use a valid test simply because there is disparate impact. Obviously, requiring SEPTA to test for the absolute "minimum" where the District Court found, and the evidence shows, that each incremental increase in an officer's aerobic capacity correspondingly increases the officer's ability to apprehend fleeing criminals and to protect fellow officers is dangerous, ludicrous, and perverts Title VII. A standard requiring the 'minimum' is particularly inappropriate since SEPTA implemented its aerobic capacity test because its officers were failing to effectively combat crime and protect its riders.

The Third Circuit's 'minimum requirements only' standard of 'business necessity' dangerously and irresponsibly requires employers to jeopardize public safety to ensure equal representation of the sexes. Its implicit social policy statement deeming the significance of the gender composition of SEPTA's police force to be more important than the safety of SEPTA's riders is reckless and repugnant to the value accorded to public safety. The Third Circuit's trade off of less public safety, for more female representation on SEPTA's police force, is particularly alarming and disturbing, in light of the unrefuted evidence that women can meet SEPTA's required aerobic capacity standard through minimal training. Clearly, SEPTA should not be compelled to quantify how many lives it is willing to endanger to ensure that it has set its cut-off score at the absolute minimum.

NOTES *(Continued)*

men is professionally acceptable with a valid test. The opinions of these testing experts, and the findings of fact of the District Court, who heard and weighed the evidence, stand in stark contrast to the Third Circuit's "minimum qualifications" standard. Further, the Third Circuit opined that the Principles for the Validation and Use of Personnel Selection Procedures ("SIOP Principles") (reproduced in part at App. E), are not to be considered, because the Court incorrectly believed that they were "inconsistent with the mission of *Griggs*." The Third Circuit did not explain how it reached this unsupported conclusion. Even the Department of Justice recognized and conceded that *Griggs* authorizes an employer to improve its work force. *See also*, 29 C.F.R. §1607.5(C) (incorporating the American Psychological Association's testing precepts, embodied in the SIOP Principles).

The Third Circuit, in footnote 21, articulated an alternative to its "minimum qualifications" only standard, allowing SEPTA to set its cut-off score at a point that excludes 80% of *both* sexes. (*See* App. A, at 23a). This footnote reveals its true agenda – equal representation of the sexes. Faced with evidence that unequivocally shows that greater levels of aerobic capacity will better protect the public, the Third Circuit willfully ignored the ability of women to train minimally before applying for this safety-sensitive position, so that they, too, can adequately protect the public and meet the demands of the job. Contrary to the Third Circuit's transparent agenda, neither Title VII, nor the Guidelines, require equal or proportionate representation of the sexes. *See e.g.*, 29 C.F.R. §1607.3(A). [19]

Title VII was never meant to reduce the quality of the work force to the lowest common denominator – yet – this is the precise consequence of the Third Circuit's "minimum qualifications" standard. The Third Circuit, in footnote 24, (App. A, at p. 25a) opines that because some incumbents have failed to meet SEPTA's aerobic capacity standard are still employed by SEPTA as transit officers, SEPTA must not have tested for the "minimum qualifications" of the transit officer position. If this logic is followed, an employer could never upgrade the performance of its workforce through raising hiring standards without first firing all of its incumbent employees who fall below the standard.

The Third Circuit inappropriately relied on SEPTA's failure to fire incumbent officers who do not meet its aerobic capacity standard, as the basis to equate the performance of

---

19. In order to fail 80% of the men and 80% of the women, SEPTA would be required to gender adjust its test, and set different passing scores for each sex, in order to achieve the same failure rates for men and women. 42 U.S.C.§2000e-2(l) arguably prohibits an employer from using different cut-off scores for men and women. *See also*, Judge Weis' dissent at App. A, at p. 45a. The fact that the Third Circuit is requiring "minimum qualifications," or alternatively, is instructing employers to equalize the failure rates for both sexes, highlights that the Third Circuit has reverted to reducing, rather than improving, productivity and job performance to accomplish, what in effect, would result in quota hiring, despite their protestations to the contrary.

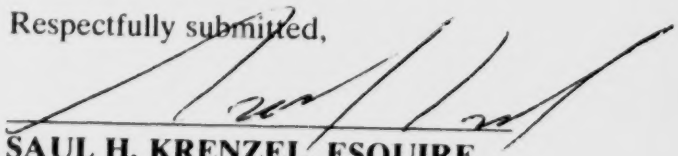
those incumbent officers with the "minimum qualifications" of the job. This flawed reasoning ignores that SEPTA is prohibited from firing or disciplining these officers, and disregards SEPTA's determination that *this* level of performance is inadequate to combat crime and was the very impetus for SEPTA's implementation of the test under review. The Third Circuit's attempt to use the performance level of incumbent officers who failed to meet SEPTA's aerobic capacity test as a proxy for the "minimum qualifications" of the job sets a dangerous precedent that completely usurps an employer's prerogative and expertise in defining the performance standards necessary to perform the job.

The Third Circuit's "minimum qualifications" standard, coupled with its view of incumbent work performance, compels the strange result than an employer can not raise the work performance of its employees, through pre-employment testing, above that of the least effective incumbent employees if disparate impact in the applicant pool is demonstrated. Surely, the Civil Rights Act of 1991 neither compels, nor sanctions, this twisted outcome. It is astonishing that the Third Circuit has required the "minimum," when in so doing it jeopardizes the safety of hundreds of thousands of SEPTA riders and the safety of fellow officers.

CONCLUSION

This Petition for Certiorari should be granted for the following reasons: (1) to clarify and define the standard of 'business necessity' to replace the varying and inconsistent ambiguous standards articulated by the Circuits; (2) to resolve the split in the Circuits as to whether safety considerations should be considered in assessing the employer's burden in demonstrating 'business necessity,' and to definitively rule on whether an employer has a lighter burden where public safety is implicated; and (3) to decide whether the Third Circuit's novel "minimum qualifications" standard of 'business necessity,' if accepted by this Court, must be equated with the inadequate performance level of incumbent employees.

Respectfully submitted,



SAUL H. KRENZEL, ESQUIRE
The Robinson Building
42 South 15th Street, Suite 800
Philadelphia, PA 19102
(215) 977-7230
Counsel for Petitioner

APPENDIX

Filed June 29, 1999

UNITED STATES COURT OF APPEALS
FOR THE THIRD CIRCUIT

Nos. 98-1644 and 98-1755

CATHERINE NATSU LANNING; ALTOVISE LOVE;
BELINDA KELLY DODSON; DENISE DOUGHERTY;
LYNNE ZIRILLI

v.

SOUTHEASTERN PENNSYLVANIA TRANSPORTATION
AUTHORITY (SEPTA);
(D.C. Civil No. 97-cv-00593)

UNITED STATES OF AMERICA

v.

SOUTHEASTERN PENNSYLVANIA TRANSPORTATION
AUTHORITY
(D.C. Civil No. 97-cv-01161)
Catherine Natsu Lanning, Altovise Love, Belinda Kelly
Dodson, Denise Doughtery and Lynne Zirilli,

Appellants in No. 98-1644

United States of America,
Appellant in No. 98-1755

Appeal from the United States District Court
for the Eastern District of Pennsylvania
(D.C. Civ. Nos. 97-cv-00593; 97-cv-01161)
District Judge: Honorable Clarence C. Newcomer

Argued: April 28, 1999

Before: MANSMANN, WEIS and GIBSON,*
Circuit Judges.

---

*Honorable John R. Gibson, of the United States Court of Appeals for the Eighth Circuit, sitting by designation.

(Filed June 29, 1999)

Lisa M. Rau, Esquire (ARGUED)
Jules Epstein, Esquire
Kairys, Rudovsky, Epstein,
Messing & Rau
924 Cherry Street, Suite 500
Philadelphia, PA 19107

Michael Churchill, Esquire
Public Interest Law Center of
Philadelphia
125 South 9th Street
Suite 700
Philadelphia, PA 19107

Counsel for Appellants: Catherine
Natsu Lanning; Altovise Love;
Belinda Kelly Dodson; Denise
Dougherty; Lynne Zirilli in
No. 98-1644

Bill Lann Lee, Esquire
Acting Assistant Attorney General
Dennis J. Dimsey, Esquire
Leslie A. Simon, Esquire
Robert S. Libman, Esquire
(ARGUED)
United States Department of Justice
Civil Rights Division
P.O. Box 66078
Washington, DC 20035-6078

Counsel for Appellant in
No. 98-1644

Saul H. Krenzel, Esquire (ARGUED)
Saul H. Krenzel & Associates
The Robinson Building, Suite 800
42 South 15th Street
Philadelphia, PA 19102

Counsel for Appellee — SEPTA

OPINION OF THE COURT

MANSMANN, *Circuit Judge.*

In this appeal, we must determine the appropriate legal standard to apply when evaluating an employer's business justification in an action challenging an employer's cutoff score on an employment screening exam as discriminatory under a disparate impact theory of liability. We hold today that under the Civil Rights Act of 1991, a discriminatory cutoff score on an entry level employment examination must be shown to measure the minimum qualifications necessary for successful performance of the job in question in order to survive a disparate impact challenge. Because we find that the District Court did not apply this standard in evaluating the employer's business justification for its discriminatory cutoff score in this case, we will reverse the District Court's judgment and remand for reconsideration under this standard. In light of our decision to remand on this basis, we need not reach the parties' other assertions of error.

I.

This appeal comes to us from a judgment entered by the District Court in favor of the Southeastern Pennsylvania Transportation Authority ("SEPTA") after a twelve day bench trial in January of 1998. Although the parties generally do not dispute the facts relevant to this appeal, to the extent there are favorable inferences to be drawn, we must draw them in favor of SEPTA as the prevailing party. In addition, because we must not disturb the factual findings of the District Court unless clearly erroneous, much of the following background is adopted from the facts as found by the District Court in its extensive memorandum opinion. *See Lanning v. Southeastern Pennsylvania Transp. Auth.*, 1998 WL 341605, at *1-*52 (E.D. Pa. June 25, 1998).

A.

SEPTA is a regional mass transit authority that operates principally in Philadelphia, Pennsylvania. In 1989, in response to a perceived need to upgrade the quality of its transit police force, SEPTA initiated an extensive program designed to improve the department. As part of this program, SEPTA dedicated its transit officers primarily to patrolling the subways and limited their responsibilities to serve as guards at other SEPTA property. In addition, SEPTA increased the number of its officers from 96 to 200 and introduced a "zone concept" for the areas they patrol.[1] SEPTA also began to consider methods by which it might upgrade the physical fitness level of its police officers.

In 1991, SEPTA hired Dr. Paul Davis to develop an appropriate physical fitness test for its police officers.[2] Dr. Davis initially met with SEPTA officials in order to ascertain SEPTA's objectives. Dr. Davis determined that SEPTA was interested in enhancing the level of fitness, physical vigor and general productivity of its police force. Once Dr. Davis had determined SEPTA's objectives, he went on a ride-along with SEPTA transit police and, over the course of two days and approximately twenty hours, rode the SEPTA trains in order to obtain a perspective on the expectations of SEPTA transit officers.

Dr. Davis next conducted a study with twenty experienced SEPTA officers, designated "subject matter experts" (SMEs), in an effort to determine what physical abilities are required to perform the job of SEPTA transit officer. From the responses Dr. Davis received in this study, he determined that running, jogging, and walking were

---

1. Under the zone concept, SEPTA designated eight separate zones covering the subway system. In a typical zone, one Lieutenant is assigned to command the zone. Two Sergeants are also assigned to the zone. Three shifts of officers per day tour the zone. Beats within the zones are assigned to the individual officers. Beats are reassigned periodically to familiarize the officers with the entire zone. Officers patrol their beats alone and on foot.

2. Dr. Davis is an expert exercise physiologist who has extensive experience in designing physical fitness employment tests for various law enforcement agencies.

important SEPTA transit officer tasks and that SEPTA officers were expected to jog almost on a daily basis.

Dr. Davis then asked the SMEs to determine what level of physical exertion was necessary to perform these tasks. The SMEs estimated that it was reasonable to expect them to run one mile in full gear in 11.78 minutes. Dr. Davis rejected this estimate as too low based upon his determination that any individual could meet this requirement. Ultimately, Dr. Davis recommended a 1.5 mile run within 12 minutes. Dr. Davis explained that completion of this run would require that an officer possess an aerobic capacity of 42.5 mL/kg/min, the aerobic capacity that Dr. Davis determined would be necessary to perform the job of SEPTA transit officer.[3]

Dr. Davis recommended that SEPTA use the 1.5 mile run as an applicant screening test. Dr. Davis understood that SEPTA officers would not be required to run 1.5 miles within 12 minutes in the course of their duties, but he nevertheless recommended this test as an accurate measure of the aerobic capacity necessary to perform the job of SEPTA transit police officer. Based upon Dr. Davis' recommendation, SEPTA adopted a physical fitness screening test for its applicants which included a 1.5 mile run within 12 minutes. Beginning in 1991, the 1.5 mile run was administered as the first component of the physical fitness test; if an applicant failed to run 1.5 miles in 12 minutes, the applicant would be disqualified from employment as a SEPTA transit officer.

It is undisputed that for the years 1991, 1993, and 1996, an average of only 12% of women applicants passed SEPTA's 1.5 mile run in comparison to the almost 60% of male applicants who passed.[4] For the years 1993 and 1996, the time period in question in this litigation, the pass rate

---

3. Dr. Davis initially decided that an aerobic capacity of 50 mL/kg/min was necessary to perform the job of SEPTA transit police officer. After determining that institution of such a high standard would have a draconian effect on women applicants, however, Dr. Davis decided that the goals of SEPTA could be satisfied by using a 42.5 mL/kg/min standard.

4. SEPTA contends that it did not seek applicants in 1992. Credited testimony was offered, however, that each of the six or seven women who took the 1.5 mile test in 1992 failed. Relying on this testimony, the District Court found that the disparate impact on women was slightly more pronounced than the 1991, 1993, and 1996 figures reflect. *See Lanning*, 1998 WL 341605 at *28.

for women was 6.7% compared to a 55.6% pass rate for men. In addition, research studies confirm that a cutoff of 12 minutes on a 1.5 mile run will have a disparately adverse impact on women.[5] SEPTA concedes that its 1.5 mile run has a disparate impact on women.

In conjunction with the implementation of its physical fitness screening test, SEPTA also began testing incumbent officers for aerobic capacity in 1991. SEPTA policy requires any officer who fails any portion of the incumbent fitness test to retest on the failed element within three months. For each portion of the physical fitness test that an incumbent officer fails, an interim goal is set for that officer.

SEPTA initially disciplined those incumbent officers who failed the fitness test. Due to protests by the incumbent officers' union, however, SEPTA discontinued its discipline policy and instead implemented an incentive program that rewarded incumbent officers for passing their interim fitness goals.

According to SEPTA's internal documents, significant percentages of incumbent officers of all ranks have failed SEPTA's physical fitness test.[6] By 1996, however, 86% of incumbent officers reached SEPTA's physical fitness standards. SEPTA has never taken any steps to determine whether incumbent officers who have failed the physical

---

5. For example, one proffered study showed that approximately 47% of men between the ages of 20 to 29 can perform a 1.5 mile run in 12 minutes where only 12% of women in the same age category can achieve this time. As noted by the District Court, testimony was offered that this study may not be entirely reliable because the women who participated in the study were predominately white women of higher socioeconomic status. Other research studies, however, were offered which show that men generally have a higher aerobic rate than women due to physiological differences between the sexes.

6. The District Court pointed to one document, for example, indicating that between July 1, 1994 and August 22, 1995, the percentage of uniformed personnel who failed the fitness test was as follows: a) Age group 20-30: 10% of all officers; b) Age group 30-40: 30% of all officers and 12% of all supervisors; c) Age group 40-50: 45% of all officers and 52% of all supervisors; d) Age group 50-60: 55% of all officers and 40% of all supervisors. *See Lanning*, 1998 WL 341605 at *31.

fitness test have adversely affected SEPTA's ability to carry out its mission.

SEPTA has promoted incumbent officers who have failed some or all of the components of the physical fitness test. SEPTA has also given special recognition, commendations, and satisfactory performance evaluations to incumbent officers who have failed the physical fitness test. SEPTA has never disciplined, terminated, removed, reassigned, suspended or demoted any transit officer for failing to perform the physical requirements of the job.

In addition, due to a clerical error, SEPTA hired a female officer in 1991 who failed the 1.5 mile run. This officer has subsequently been "decorated" by SEPTA and has been nominated repeatedly for awards such as Officer of the Year and Officer of the Quarter. SEPTA has commended her for her outstanding performance as a police officer and has chosen her to serve as one of SEPTA's two defensive tactics instructors.

SEPTA employs an extremely low number of women in its transit police force. The District Court found that, as of July 1997, SEPTA employed only 16 women in its 234 member police force. Only two of these women hold ranks higher than that of patrol officer. *See Lanning*, 1998 WL 341605 at *27.

B.

On January 28, 1997, after satisfying all administrative prerequisites, five women who failed SEPTA's 1.5 mile run brought a Title VII class action against SEPTA on behalf of all 1993 female applicants, 1996 female applicants and future female applicants for employment as SEPTA police officers who have been or will be denied employment by reason of their inability to meet the physical entrance requirement of running 1.5 miles in 12 minutes or less. On February 18, 1997, the Department of Justice, after conducting the appropriate investigation of SEPTA's employment practices and meeting all conditions precedent under Title VII, also filed suit on behalf of the United States challenging SEPTA's entire physical fitness test, including the 1.5 mile run. The District Court properly exercised

jurisdiction over these Title VII actions challenging SEPTA's hiring practices pursuant to 28 U.S.C. § 1331. On April 21, 1997, the District Court consolidated the two actions for all purposes up to and including trial.

After litigation commenced, SEPTA hired expert statisticians to submit reports examining the statistical relationship between the aerobic capacity of SEPTA's officers and their number of arrests, "arrest rates"[7] and number of commendations. In these reports, the statisticians concluded that there was a statistically significant correlation between high aerobic capacity and arrests, arrest rates and commendations. In addition, one expert prepared a report that estimated that 51.9% of the persons arrested for serious crimes between 1991 and 1996 had an aerobic capacity of 48 mL/kg/min and 27% of those arrested had an aerobic capacity of less than 42 mL/kg/min.[8] Based upon these reports, the District Court held that SEPTA established that its aerobic capacity requirement is job related and consistent with business necessity. *See Lanning*, 1998 WL 341605 at *35.

The District Court also found support for this conclusion in an expert report submitted on behalf of SEPTA by Dr. Robert Moffatt. Dr. Moffatt simulated a training course and concluded that officers with aerobic capacities of 45 mL/kg/min or better had a 7-8% decrement in their ability to perform physical activities after a run of approximately three minutes; officers with an aerobic capacity of less than 45 mL/kg/min exhibited a 30% decrement in physical ability after the same run. The District Court found that Dr. Moffatt's study demonstrates "the manifest relationship of aerobic capacity to the critical and important duties of a SEPTA transit police officer . . . ." *Id.* at *68.

---

7. "Arrest rates" were tabulated by expressing the number of arrests made by an officer as a percentage of the number of incident reports involving that officer. *See* App. at 3040-41 (Siskin Expert Report).

8. The category of "serious crimes" includes homicide, rape, robbery, aggravated assault, burglary, theft, and auto theft. This category of arrests accounts for approximately ten percent of all reported incidents and seven percent of all reported arrests. *See* App. at 3040. (Siskin Expert Report).

The District Court entered judgment in favor of SEPTA on all claims. Both the individual plaintiffs and the United States have taken appeals from the District Court's final judgment, over which we have jurisdiction pursuant to 28 U.S.C. § 1291. On appeal, the individual plaintiffs assert that the District Court applied incorrect legal standards in evaluating SEPTA's business necessity defense and that the District Court made erroneous findings of fact in determining that SEPTA's 1.5 mile run does not violate Title VII. Although the United States initially challenged SEPTA's implementation of its entire physical fitness test, on appeal the United States joins the individual plaintiffs in asserting error solely with respect to the District Court's determination that SEPTA's 1.5 mile run is not violative of Title VII. Because the issue of whether the District Court applied the correct legal standard is one of law, our review is plenary.

II.

Under Title VII's disparate impact theory of liability, plaintiffs establish a prima facie case of disparate impact by demonstrating that application of a facially neutral standard has resulted in a significantly discriminatory hiring pattern. *See Dothard v. Rawlinson*, 433 U.S. 321, 329 (1977). Once the plaintiffs have established a prima facie case, the burden shifts to the employer to show that the employment practice is "job related for the position in question and consistent with business necessity . . . ." 42 U.S.C. § 2000e-2k. Should the employer meet this burden, the plaintiffs may still prevail if they can show that an alternative employment practice has a less disparate impact and would also serve the employer's legitimate business interest. *See Albemarle Paper Co. v. Moody*, 422 U.S. 405, 425 (1975).

Because SEPTA concedes that its 1.5 mile run has a disparate impact on women, the first prong of the disparate impact analysis is not at issue in this appeal.[9] Rather, this

9. On appeal, SEPTA offered evidence to establish that the individual female applicants who failed SEPTA's 1.5 mile run demonstrated a cavalier attitude in preparing for and taking the test. As aptly noted by plaintiffs' counsel at oral argument, this evidence has no bearing upon our analysis in this appeal because SEPTA has conceded that its test has a severe disparate impact on women.

appeal focuses our attention on the proper standard for evaluating whether SEPTA's 1.5 mile run is "job related for the position in question and consistent with business necessity" under the Civil Rights Act of 1991. Because the Act instructs that this standard incorporates only selected segments of prior Supreme Court jurisprudence on the business necessity doctrine, we examine the history of this doctrine in order to resolve this threshold issue.

A.

The disparate impact theory of discrimination under Title VII was judicially created in the seminal case of *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971). In embracing disparate impact, the Court recognized that Title VII was meant not only to proscribe overt discrimination, but also to prohibit "practices that are fair in form, but discriminatory in operation." *Griggs*, 401 U.S. at 431. The Court made clear that what is required by Title VII is "the removal of artificial, arbitrary, and unnecessary barriers to employment when the barriers operate invidiously to discriminate on the basis of racial or other impermissible classification." *Id.* Accordingly, the Court announced that in evaluating practices fair in form but discriminatory in operation, "[t]he touchstone is business necessity." *Id.*

The Court, however, was unclear in articulating what an employer must show to demonstrate business necessity. The Court couched the employer's burden in terms of showing that its practice is "related to job performance"; "bear[s] a demonstrable relationship to successful performance of the jobs for which it was used"; has "a manifest relationship to the employment in question"; and is "demonstrably a reasonable measure of job performance." *Id.* at 431, 432, 436. In applying this standard, however, the Court rejected the employer's justification in *Griggs* that its standardized intelligence tests and diploma requirements generally would improve the overall quality of the work force in its power plant. The Court held that, although these requirements may be useful, they could not be used to exclude disproportionately a protected group when the employer failed to show that they do not test an

applicant's ability to perform the job in question. *Id.* at 431-33.

The Court next spoke to the issue of business necessity in *Albemarle Paper Co. v. Moody*, 422 U.S. 405 (1975). In *Albemarle*, an employer sought to justify the use of verbal exam and high school diploma requirements in determining whether to promote employees to more skilled positions in its paper mill. *Albemarle*, 422 U.S. at 408-11. In preparation for trial, the employer hired an industrial psychologist to complete validation studies showing that the tests were job related because they had a statistically significant correlation with supervisorial ratings in several groups of the jobs in question. *Id.* at 429-30. The Court, nevertheless, rejected the employer's contention that its requirements were job related.

The Court held that "discriminatory tests are impermissible unless shown, by professionally acceptable methods, to be 'predicative of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated.'" *Id.* at 431 (quoting 29 CFR § 1607.4(c)). In so holding, the Court noted that the Equal Employment Opportunity Commission (EEOC) Guidelines for professional standards of test validation are entitled to great deference in determining whether an employer has demonstrated that its requirements are job related. *Id.* at 430-31. The Court rejected the employer's validation studies as inadequate in several respects under the EEOC Guidelines. For example, the Court rejected the studies because they focused on the most qualified employees near the top of the line of progression, stating:

> The fact that the best of those employees working near the top of a line of progression score well on a test does not necessarily mean that that test, or some particular cutoff score on the test, is a permissible measure of the minimal qualifications of new workers entering lower level jobs.

Id. at 434. The Court accordingly held that consideration must be given to the possible use of testing as a promotion device rather than as a screen for entry into lower level

jobs. *Id.* Due to several inadequacies of the employer's validation studies, the Court held that the employer had failed to show that its requirements were job related to the position in question. *Id.* at 435-36.

The next Title VII case to raise the business necessity issue for the Court's consideration was *Dothard v. Rawlinson*, 433 U.S. 321 (1977).[10] In *Dothard*, female applicants challenged a prison's minimum height and weight requirements for its prison guard positions as violative of Title VII. On the issue of business necessity, the Court made clear that "a discriminatory employment practice must be shown to be necessary to safe and efficient job performance to survive a Title VII challenge." *Dothard*, 433 U.S. at 332 n.14. The Court rejected the prison's assertion that height and weight requirements have a relationship to the unspecified amount of strength essential to effective job performance, holding that if strength is a bona fide job related quality, the prison could test for it directly by adopting and validating a fairly administered strength test. *Id.* at 331-32.

The Court's next definitive statement on the business necessity doctrine is found in *Wards Cove Packing Co., Inc. v. Atonio*, 490 U.S. 642 (1989), where a majority of the Court deviated from its previous business necessity jurisprudence in adopting a more liberal test for business necessity.[11] According to the Court:

---

10. Prior to *Dothard*, the Court included some language related to the business necessity doctrine in *Washington v. Davis*, 426 U.S. 229 (1976), an equal protection case. Because *Washington* is not a Title VII case, however, we cannot treat the language in *Washington* as reflective of the pre-*Wards Cove* business necessity doctrine applicable to Title VII cases.

11. Two cases prior to *Wards Cove* forecast some of the changes to come. In *New York City Transit Auth. v. Beazer*, 440 U.S. 568 (1979), the Court disposed of a Title VII case by holding that the plaintiffs failed to establish a prima facie case of disparate impact. The Court, however, commented on the business necessity doctrine in dicta. In a footnote, the Court stated that even if a prima facie case had been established, the employer would have shown business necessity by establishing that its practice significantly serves its legitimate business goals of safety and efficiency. *Beazer*, 440 U.S. at 587 n.31. Similarly, a plurality opinion in

> [T]he dispositive issue is whether a challenged practice serves, in a significant way, the legitimate employment goals of the employer. The touchstone of this inquiry is a reasoned review of the employer's justification for his use of the challenged practice. A mere insubstantial justification in this regard will not suffice, because such a low standard of review would permit discrimination to be practiced through the use of spurious, seemingly neutral employment practices. At the same time, though, there is no requirement that the challenged practice be "essential" or "indispensable" to the employer's business for it to pass muster . . . .

Wards Cove, 490 U.S. at 659 (citations omitted). In addition, the Court made clear that at the business necessity stage of Title VII litigation, the employer bears only the burden of production; the burden of persuasion remains on the disparate impact plaintiff at all times. *Id.* As we have previously recognized, the *Wards Cove* standard may reasonably be viewed as a departure from the more stringent business necessity standard under *Griggs* and its progeny. *See Newark Branch, N.A.A.C.P. v. Town of Harrison, New Jersey*, 940 F.2d 792, 803 (3d Cir. 1991)(noting that *Wards Cove* "arguably diluted the business necessity burden" under *Griggs*).

B.

In response to *Wards Cove*, Congress enacted the Civil Rights Act of 1991. One of the primary purposes of the Act was "to codify the concepts of 'business necessity' and 'job related' enunciated by the Supreme Court in *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971), and in the other Supreme

---

Watson v. Fort Worth Bank & Trust, 487 U.S. 977 (1988), suggested that employers could meet their burden of establishing business necessity simply by advancing a legitimate business reason for the practice in question. *Watson*, 487 U.S. at 998. While the language in these cases clearly foreshadowed the Court's holding in *Wards Cove*, this language had never been embraced by a majority of the Court as the binding standard for business necessity prior to *Wards Cove*.

Court decisions prior to *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642 (1989)." Civil Rights Act of 1991, Pub L. No. 102-166, § 3, 105 Stat. 1071, 1071 (1992). As part of this codification of *Griggs*, the Act made clear that both the burden of production and the burden of persuasion in establishing business necessity rest with the employer. *See* 42 U.S.C. § 2000e-2(k).

In addition, the Act codified the business necessity doctrine by using the following language:

> An unlawful employment practice based on disparate impact is established under this subchapter only if —
>
> (i) a complaining party demonstrates that a respondent uses a particular employment practice that causes a disparate impact on the basis of race, color, religion, sex, or national origin and the respondent fails to demonstrate that the challenged practice is *job related for the position in question and consistent with business necessity*; or
>
> (ii) the complaining party makes the demonstration described in subparagraph (C) with respect to an alternative employment practice and the respondent refuses to adopt such alternative employment practice.

42 U.S.C. § 2000e-2(k)(1)(A)(emphasis added). The Act further instructs that in interpreting its business necessity language, "[n]o statements other than the interpretive memorandum . . . shall be considered legislative history of, or relied upon in any way as legislative history . . . ." Civil Rights Act of 1991, Pub L. No. 102-166, § 105(b), 105 Stat. 1071, 1075 (1992). The interpretive memorandum referenced in this portion of the Act states in relevant part:

> The terms "business necessity" and "job related" are intended to reflect the concepts enunciated by the Supreme Court in *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971), and in the other Supreme Court decisions prior to *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642 (1989).

137 Cong. Rec. 28,680 (1991). After the passage of the Act, proponents of both a strict test for business necessity and

a more liberal requirement claimed victory in the standard adopted by the Act.[12]

III.

The Supreme Court has yet to interpret the "job related for the position in question and consistent with business necessity" standard adopted by the Act. In addition, our sister courts of appeals that have applied the Act's standard to a Title VII challenge have done so with little analysis. *See, e.g., Fitzpatrick v. City of Atlanta*, 2 F.3d 1112, 1117-18 (11th Cir. 1993)(noting that Civil Rights Act of 1991 statutorily reversed *Wards Cove* but ruling in favor of employer because practice was demonstrably necessary to meet an "important business goal"); *Bradley v. Pizzaco of Nebraska, Inc.*, 7 F.3d 795, 797-98 (8th Cir. 1993)(noting that *Griggs* standard was reinstated by the Act and holding that employer failed to meet *Griggs* standard).

Because the Act proscribes resort to legislative history with the exception of one short interpretive memorandum endorsing selective caselaw, our starting point in interpreting the Act's business necessity language must be that interpretive memorandum. The memorandum makes clear that Congress intended to endorse the business necessity standard enunciated in *Griggs* and not the *Wards Cove* interpretation of that standard. By Congress' distinguishing between *Griggs* and *Wards Cove*, we must conclude that Congress viewed *Wards Cove* as a significant departure from *Griggs*. Accordingly, because the Act clearly chooses *Griggs* over *Wards Cove*, the Court's interpretation

---

12. *See* Andrew C. Spiropoulos, *Defining the Business Necessity Defense to the Disparate Impact Cause of Action: Finding the Golden Mean*, 74 N.C. L. Rev. 1479, 1516-20 (1996)(outlining the respective positions of both sides to the debate); *compare also* Michael Carvin, *Disparate Impact Claims Under the New Title VII*, 68 Notre Dame L. Rev. 1153 (1993)(arguing that *Wards Cove* is still good law after Civil Rights Act of 1991); *with* Susan S. Grover, *The Business Necessity Defense in Disparate Impact Discrimination Cases*, 30 Ga. L. Rev. 387 (1996)(arguing for a strict business necessity standard under the Act); Note, *The Civil Rights Act of 1991: The Business Necessity Standard*, 106 Harv. L. Rev. 896 (1993)(asserting that *Wards Cove* does not survive the Act).

of the business necessity standard in *Wards Cove* does not survive the Act.[13]

We turn now to articulate the standard for business necessity - one most consistent with *Griggs* and its pre-*Wards Cove* progeny. The laudable mission begun by the Court in *Griggs* was the eradication of discrimination through the application of practices fair in form but discriminatory in practice by eliminating *unnecessary* barriers to employment opportunities. In the context of a hiring exam with a cutoff score shown to have a discriminatory effect, the standard that best effectuates this mission is implicit in the Court's application of the business necessity doctrine to the employer in *Griggs*, *i.e.*, that a discriminatory cutoff score is impermissible unless shown to measure the minimum qualifications necessary for successful performance of the job in question. Only this standard can effectuate the mission begun by the Court in *Griggs*; only by requiring employers to demonstrate that their discriminatory cutoff score measures the minimum qualifications necessary for successful performance of the job in question can we be certain to eliminate the use of excessive cutoff scores that have a disparate impact on minorities as a method of imposing unnecessary barriers to employment opportunities.

The evolution of the Court's articulation of the business necessity doctrine in both *Albemarle* and *Dothard* reinforces

---

13. We are cognizant that a contrary argument has been advanced in which it is asserted that *Wards Cove* remains the controlling standard. *See* Carvin, *supra* note 12, at 1157-64. Pursuant to the argument, the business necessity standard announced in *Wards Cove* simply clarified *Griggs* and therefore is not inconsistent with the Act's command to apply the standard enunciated in *Griggs*. In addition, it is asserted that due to the legislative history of the Act, it would be improper to apply a strict business necessity standard. This argument, however, ignores two important aspects of the Act which constrain our interpretation of the standard adopted. First, the interpretive memorandum's distinction between *Griggs* and *Wards Cove* casts significant doubt on the assertion that Congress read *Wards Cove* as simply a clarification of *Griggs*. Second, the Act precludes us from considering the legislative history upon which this argument relies for support. Accordingly, we find this argument to be devoid of merit.

the conclusion that this standard is both implicit in *Griggs* and central to its mission. In *Albemarle*, the Court explained that discriminatory tests must be validated to show that they are "predictive of . . . important elements of work behavior which comprise . . . the job . . . for which candidates are being evaluated" and that the scores of the higher level employees do not necessarily validate a cutoff score for the minimum qualifications to perform the job at an entry level. *Albemarle*, 422 U.S. at 431, 434. This is simply another way of saying that discriminatory cutoff scores must be validated to show they measure the minimum qualifications necessary for successful performance of the job. Similarly, in *Dothard*, the Court made clear that "a discriminatory employment practice," such as a discriminatory cutoff score on an entry level exam, "must be shown to be necessary to safe and efficient job performance to survive a Title VII challenge." *Dothard*, 433 U.S. at 332 n.14.

Taken together, *Griggs*, *Albemarle* and *Dothard* teach that in order to show the business necessity of a discriminatory cutoff score an employer must demonstrate that its cutoff measures the minimum qualifications necessary for successful performance of the job in question. Furthermore, because the Act instructs us to interpret its business necessity language in conformance with *Griggs* and its pre-*Wards Cove* progeny, we must conclude that the Act's business necessity language incorporates this standard.

Our conclusion that the Act incorporates this standard is further supported by the business necessity language adopted by the Act. Congress chose the terms "job related for the position in question" *and* "consistent with business necessity." Judicial application of a standard focusing solely on whether the qualities measured by an entry level exam bear some relationship to the job in question would impermissibly write out the business necessity prong of the Act's chosen standard. With respect to a discriminatory cutoff score, the business necessity prong must be read to demand an inquiry into whether the score reflects the minimum qualifications necessary to perform successfully the job in question. *See also* EEOC Guidelines, 29 C.F.R. §1607.5 (H) (noting that cutoff scores should "be set so as

to be reasonable and consistent with normal expectations of acceptable proficiency within the work force.").

In addition, Congress' decision to emphasize the importance of the policies underlying the disparate impact theory of discrimination through its codification supports application of this standard to discriminatory cutoff scores. The disparate impact theory of discrimination combats not intentional, obvious discriminatory policies, but a type of covert discrimination in which facially neutral practices are employed to exclude, unnecessarily and disparately, protected groups from employment opportunities. Inherent in the adoption of this theory of discrimination is the recognition that an employer's job requirements may incorporate societal standards based not upon necessity but rather upon historical, discriminatory biases.[14] A business necessity standard that wholly defers to an employer's judgment as to what is desirable in an employee therefore is completely inadequate in combating covert discrimination based upon societal prejudices. Only a business necessity doctrine that examines discriminatory cutoff scores in light of the minimum qualifications that are necessary to perform the job in question successfully can address adequately this subtle form of discrimination.[15]

---

14. For an interesting discussion on male-oriented biases in the labor market see Maxine N. Eichner, *Getting Women Work That Isn't Women's Work: Challenging Gender Biases in the Workplace Under Title VII*, 97 Yale L.J. 1397 (1988). *See also, Hurley v. The Atlantic City Police Dept.*, ___ F.3d. ___, ___ n.5, 1999 WL 150301 (3d Cir. 1999)(noting egregious sexual harassment to which a female police officer was subjected by her male colleagues); *Mazus v. Department of Transp., Com. of Pa.*, 629 F.2d 870, 876 (3d Cir. 1980)(Sloviter, J., dissenting)(noting allegations demonstrating prevalent male attitude that construction work is not the "type of work" women should perform).

15. We need not be concerned that implementation of this standard will result in forcing employers to adopt quotas, a result that would be inconsistent with the mandates of Title VII. If an employer can demonstrate that its discriminatory cutoff score reflects the minimum qualifications necessary for successful job performance, it will be able to continue to use it. If not, the employer must abandon that cutoff score, but is free to develop either a non-discriminatory practice which furthers its goals, or an equally discriminatory practice that can meet this

Accordingly, we hold that the business necessity standard adopted by the Act must be interpreted in accordance with the standards articulated by the Supreme Court in *Griggs* and its pre-*Wards Cove* progeny which demand that a discriminatory cutoff score be shown to measure the minimum qualifications necessary for the successful performance of the job in question in order to survive a disparate impact challenge.[16]

---

standard. Nothing in the *Griggs* business necessity standard requires employers to hire employees in numbers to reflect the ethnic, racial or gender make-up of the community.

The following example based upon the facts of this case illustrates this point. Assuming that SEPTA's 1.5 mile run has a disparate impact on women and that SEPTA can not show that the 12 minute cutoff measures the minimum aerobic capacity necessary to be a successful transit officer, it does not follow that SEPTA would then be required to hire women in equal proportion to men. Several options would be available to SEPTA. For example, SEPTA could: 1) abandon the test as a hiring requirement but maintain an incentive program to encourage an increase in the officers' aerobic capacities; 2) validate a cutoff score for aerobic capacity that measures the minimum capacity necessary to successfully perform the job and maintain incentive programs to achieve even higher aerobic levels; or 3) institute a non-discriminatory test for excessive levels of aerobic capacity such as a test that would exclude 80% of men as well as 80% of women through separate aerobic capacity cutoffs for the different sexes. Each of these options would help SEPTA achieve its stated goal of increasing aerobic capacity without running afoul of Title VII and none of these options require hiring by quota.

16. Relying upon *Spurlock v. United Airlines, Inc.*, 475 F.2d 216 (10th Cir. 1972), and like cases from our sister courts of appeals, the dissent asserts that this standard should not apply to SEPTA because the job of SEPTA transit officer implicates issues of public safety. Under the Act, however, our interpretation of the business necessity language is limited to "the concepts enunciated by the Supreme Court in *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971), and in the other *Supreme Court* decisions prior to *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642 (1989)." *See* 137 Cong. Rec. 28,680 (1991)(emphasis added). Because the Supreme Court never adopted the holding of *Spurlock* prior to *Wards Cove*, its is clear that, under the Act, we are not to consider *Spurlock* as authoritative. Furthermore, if Congress had intended to endorse the holding of *Spurlock*, it could have done so affirmatively. Accordingly, because the Act limits our interpretation to Supreme Court

IV.

Although the District Court purported to apply the Act's "job related to the position in question and consistent with business necessity" standard to SEPTA's cutoff score on its 1.5 mile run, it is clear from the District Court's memorandum opinion that it did not apply the standard we have found to be implicit in *Griggs* and incorporated by the Act. The District Court rejected the formulation of the *Griggs* standard found in *Dothard*, characterizing it as dicta, and relied instead upon language found in *New York City Transit Auth. v. Beazer*, 440 U.S. 568 (1979). As our prior discussion makes clear,[17] the *Beazer* language is dicta and the *Dothard* standard is binding under the Act. Moreover, the *Beazer* dicta upon which the District Court relied mirrors the standard adopted by *Wards Cove*. *Compare Lanning*, 1998 WL 341605 at *54 (noting that in *Beazer*, the Court "implicitly approves employment practices that significantly serve, but are neither required by nor necessary to, the employer's legitimate business interests") *with Wards Cove*, 490 U.S. at 659 (stating that

---

jurisprudence and does not otherwise endorse *Spurlock*, we are not at liberty to adopt the holding of *Spurlock* at this juncture. Moreover, to the extent that *Spurlock* and other cases from our sister courts of appeals can be read to suggest that minimum qualifications do not apply to certain types of employment, these cases are inconsistent with the teachings of *Griggs* and are accordingly uninformative under the Act.

Furthermore, to the limited extent that the Supreme Court's pre-*Wards Cove* jurisprudence instructs that public safety is a legitimate consideration, application of the business necessity standard to SEPTA is consistent with that jurisprudence because the standard itself takes public safety into consideration. If, for example, SEPTA can show on remand that the inability of a SEPTA transit officer to meet a certain aerobic level would significantly jeopardize public safety, this showing would be relevant to determine if that level is necessary for the successful performance of the job. Clearly a SEPTA officer who poses a significant risk to public safety could not be considered to be performing his job successfully. We are accordingly confident that application of the business necessity standard to SEPTA is fully consistent with the Supreme Court's pre-*Wards Cove* jurisprudence as required by the Act.

17. *See supra* note 11.

standard is "whether a challenged practice serves, in a significant way, the legitimate employment goals of the employer" and noting that there is no requirement that the practice be essential). As we previously stated, the *Wards* Cove standard does not survive the Act.

The District Court's application of its understanding of business necessity to SEPTA's business justification further illustrates that the District Court did not apply the correct legal standard. As an initial matter, the District Court seemed to conclude that Dr. Davis' expertise alone is sufficient to justify the 42.5 mL/kg/min aerobic capacity cutoff measured by the 1.5 mile run.[18] This conclusion disregards the teachings of *Griggs*, *Albemarle* and *Dothard* in which the Court made clear that judgment alone is

---

18. While relying predominately upon Dr. Davis' expertise, the District Court does point to a study which Dr. Davis completed for Anne Arundel County, Maryland in which he concluded that a 42.5 mL/kg/min aerobic capacity predicted success as an Anne Arundel County police officer. Absent a finding that the work of an Anne Arundel County police officer is comparable to SEPTA transit officer work, a finding the District Court did not make, reliance on this validation study is misplaced. *See* 29 C.F.R. § 1607.7 (B)(2); *see also* 29 C.F.R. § 1607.7(B)(3)(explaining that validation studies created for other employers must also include a study of "test fairness"). Furthermore, it is unclear from Dr. Davis' report whether the Anne Arundel study's 42.5 mL/kg/min cutoff actually measures for qualities significant to SEPTA transit police performance. *Compare* App. at 3134 (Davis Report) (noting that 42.5 mL/kg/min level for Anne Arundel study is significant for carrying an unspecified amount of weight and generally effecting arrests) *with* App. at 3132 (Davis Report) (stating "[t]ransit police officers are more likely to have incidents come to them, as opposed to responding to the scene of an event. By mission, the presence of the officer is that of a deterrent, maintaining maximum visibility. Occasionally, officers will come upon criminal activities such as assaults or robberies, but for the most part, the officer will attempt to control a situation such as disorderly conduct or force compliance (paying fares) without having to make an arrest."); *see also* App. at 3139 (Davis Report)(quoting experienced officer as stating "[t]he most important factors in my opinion of being a good officer is to be able to think clearly at all times an [sic] verbalize and or articulate when dealing with all people . . . . Running quickly is physically demanding, although in the transit system, most dealings are close, physical altercations."). In addition, it is unclear from the record whether the Anne Arundel study itself was properly validated.

insufficient to validate an employer's discriminatory practices.[19] More fundamentally, however, nowhere in its extensive opinion did the District Court consider whether Dr. Davis' 42.5 mL/kg/min cutoff reflects the minimum aerobic capacity necessary to perform successfully the job of SEPTA transit police officer.

Instead, the District Court upheld this cutoff because it was "readily justifiable." *Lanning*, 1998 WL 341605 at *57.[20] The validation studies of SEPTA's experts upon which the District Court relied to support this conclusion demonstrate the extent to which this standard is insufficient under the Act. The general import of these studies is that the higher an officer's aerobic capacity, the better the officer is able to perform the job. Setting aside the validity of these studies, this conclusion alone does not validate Dr. Davis' 42.5 mL/kg/min cutoff under the Act's business necessity standard.[21] At best, these studies show that aerobic

(Text continued on page 24)

---

19. The danger of allowing an employer to carry its burden by relying simply upon an expert's unvalidated judgment as to an appropriate cutoff score in a testing device is illustrated by this case. In determining an appropriate cutoff for aerobic capacity, Dr. Davis rejected the SMEs' estimate of the minimal qualifications necessary to perform the job even though these SMEs were experienced transit officers. Dr. Davis then determined that "a SEPTA transit officer *needs* an aerobic capacity of 50 mL/kg/min to successfully perform a number of tasks." *Lanning*, 1998 WL 341605 at *16 (emphasis added). Dr. Davis, however, revised this requirement, finding that "the goals of SEPTA could be satisfied by using a 42.5 mL/kg/min standard" after determining that the higher limit would have a "draconian" effect on women. *Id.* There is no indication in the District Court's opinion as to how Dr. Davis determined that the lower standard would be sufficient. Where, as here, the cutoff score chosen has a discriminatory disparate impact, *Griggs* prohibits the establishment of exactly this type of arbitrary barrier to employment opportunities.

20. The District Court seems to have derived this standard from the *Principles for the Validation and Use of Personnel Selection Procedures* ("SIOP Principle"), principles published by the Society for Industrial and Organizational Psychology as a professional guideline for conducting validation research and personnel selection. To the extent that the SIOP Principles are inconsistent with the mission of *Griggs* and the business necessity standard adopted by the Act, they are not instructive.

21. The Court has cautioned that studies done in anticipation of litigation to validate discriminatory employment tests that have already

been given must be examined with great care due to the danger of lack of objectivity. *Albemarle*, 422 U.S. at 433 n.32. We also have warned in a disparate impact context that "the story statistics tell depends, not unlike beauty, upon the eye and ear of the beholder" and that "we must apply a critical and cautious ear to one dimensional statistical presentation." *Bryant v. International Sch. Servs., Inc.*, 675 F.2d 562, 573 (3d Cir. 1982). A critical evaluation of the statistical studies relied upon by the District Court in this case, reveals several aspects of these studies that we find to be, at a minimum, disconcerting.

The following concerns are only a representative sample of possible deficiencies in these studies: 1) While the ability to make an arrest may be an important aspect of the job, the absolute number of arrests or "arrest rates" do not necessarily correlate with successful job performance. *See* App. at 3132 (noting that SEPTA officer should generally attempt to control a situation without having to make an arrest); 2) The study on arrests and arrest rates examined a disproportionately large number of officers with an aerobic capacity over 42 mL/kg/min compared to the number of officers with an aerobic capacity under that level which likely skewed the results. *See, e.g.*, App. at 3053 (comparing arrests of 231 officers with aerobic capacities under the 42 mL/kg/min with arrests of 813 officers with aerobic capacities over the 42 mL/kg/min); *see also*, 29 C.F.R. § 1607.14(B)(6)(noting that "[r]eliance upon a selection procedure which is significantly related to a criterion measure, but which is based upon a study involving a large number of subjects and has a low correlation coefficient will be subject to close review if it has a large adverse impact."); 3) The comparison of aerobic capacity with commendations is not helpful absent finding as to the subjective considerations involved in awarding commendations. *See Albemarle*, 422 U.S. at 432-33; 4) The studies' emphasis on arrests for "serious crimes" is suspect; these arrests account for only 7% of all arrests and therefore represent only a small aspect of job. *See generally* 29 C.F.R. § 1607.14(B)(6)(noting that reliance on single selection instrument which is related to only one of many job duties will be subject to close review); 5) SEPTA's table on the field performance of its officers belies the contention that there is a strict linear relationship of arrests to aerobic capacity; officers at less than 37 mL/kg/min had an average arrests of 13.6 compared to officers with at least a 48 mL/kg/min level who had average arrests of 13.9. *See* App. at 3065 (Defendant's Exhibit 52D); 6) The study on the average aerobic capacity of perpetrators has little meaning unless SEPTA can show that arrests of these perpetrators are typically aerobic contests; because SEPTA police are armed, such a showing is unlikely.

capacity is related to the job of SEPTA transit officer. A study showing that "more is better," however, has no bearing on the appropriate cutoff to reflect the minimal qualifications necessary to perform successfully the job in question.

Dr. Siskin's testimony is particularly instructive on this point. Dr. Siskin testified that in view of the linear relationship between aerobic capacity and the arrest parameters, any cutoff score can be justified since higher aerobic capacity levels will get you more field performance (*i.e.*, "more is better"). *See Lanning*, 1998 WL 341605 at *41. Under the District Court's understanding of business necessity, which requires only that a cutoff score be "readily justifiable," SEPTA, as well as any other employer whose jobs entail any level of physical capability, could employ an unnecessarily high cutoff score on its physical abilities entrance exam in an effort to exclude virtually all women by justifying this facially neutral yet discriminatory practice on the theory that more is better.[22] This result contravenes *Griggs* and demonstrates why, under *Griggs*, a discriminatory cutoff score must be shown to measure the

---

Because we are remanding for the District Court to reconsider this evidence in light of the *Griggs* standard, we need not rule on whether any of the District Court's prior findings as to these studies were clearly erroneous. We comment here on the validity of these studies only to draw the District Court's attention to these concerns and to encourage the District Court to take a critical look at these studies, if necessary, on remand.

22. Such a result has the potential to have a significant detrimental impact on the amount and type of employment opportunities available to women. Obviously, under a "more is better" theory, employers such as police departments, fire departments and correctional facilities could develop physical tests with unnecessarily high cutoffs that would effectively exclude women from their ranks. Perhaps less obvious, however, is the impact that this result could have on industries where strength even minimally related to the job in question. For example, all companies engaged in delivery, construction or any other type of physical labor would be permitted to develop unnecessary strength requirements on the theory that "more is better" or "the stronger the worker, the faster the job gets done." This result is clearly unacceptable given the policies underlying both Title VII and the disparate impact theory of discrimination.

minimum qualifications necessary to perform successfully the job in question.[23]

V.

For the foregoing reasons, it is clear to us that the District Court did not employ the business necessity standard implicit in *Griggs* and incorporated by the Act which requires that a discriminatory cutoff score be shown to measure the minimum qualifications necessary for successful performance of the job in question in order to survive a disparate impact challenge. We will therefore vacate the judgment of the District Court and remand this appeal for the District Court to determine whether SEPTA has carried its burden of establishing that its 1.5 mile run measures the minimum aerobic capacity necessary to perform successfully the job of SEPTA transit police officer.[24] Because this is the first occasion we have had to clarify the

---

23. This is not to say that studies that actually prove that "more is better" are always irrelevant to validation of an employer's discriminatory practice. For example, a content validated exam, such as a typing exam for the position of typist, which demonstrates that the applicants who score higher on the exam will exhibit better job performance may justify a rank-ordering hiring practice that is discriminatory. In such a case, a validation study proving that "more is better" may suffice to validate the rank-order hiring. This is true, however, in only the rarest of cases where the exam tests for qualities that fairly represent the totality of a job's responsibilities. It is unlikely that such a study could validate rank-hiring with a discriminatory impact based upon physical attributes in complex jobs such as that of police officer in which qualities such as intelligence, judgment, and experience surely play a critical role. This is especially true in SEPTA's case, where the record indicates that SEPTA patrol officers encounter "running assists," the most strenuous task upon which SEPTA's aerobic capacity testing predominately was justified, at an average rate of only twice per year. *Compare Lanning*, 1998 WL 341605 at *5 (finding that SEPTA has approximately 380 running assists per year) *with id.* at *27 (noting that SEPTA has 190 patrol officers).

24. The District Court rejected as irrelevant the plaintiffs' evidence that incumbent officers had failed the physical fitness test yet successfully performed the job and that other police forces function well without an aerobic capacity admission test. *See Lanning*, 1998 WL 341605 at *68-*70. Under the standard implicit in *Griggs* and incorporated into the Act, this evidence tends to show that SEPTA's cutoff score for aerobic capacity does not correlate with the minimum qualifications necessary to perform successfully the job of SEPTA transit officer. Accordingly, this evidence is relevant and should be considered by the District Court on remand.

Act's business necessity standard, on remand the District Court may wish to exercise its discretion to allow the parties to develop further the record in keeping with the standard announced here.

WEIS, *Circuit Judge, dissenting*:

The "minimum qualifications" criterion of business justification does not apply to all types of employment. When public safety is at stake, a lighter burden is placed on employers to justify their hiring requirements. Because I believe that the latter standard applies in this case, I would affirm.

I.

Concerned about its inability to control crime on its property, SEPTA instituted a three-pronged attack on the problem. It added a substantial number of officers, implemented a zone method of patrol, and adopted standards to improve the generally poor physical condition of its officers. Unlike many metropolitan police departments, SEPTA officers are deployed alone and on foot, engaging in physical activities more frequently than other law enforcement agencies.

The patrol zones present significant variations in conditions that affect the physical exertion of officers in the performance of their duties. Zone One, for example, has a climb of 30 to 50 steps from street level. Zone Three, a mixture of above- and below-ground locations, borders a large shopping mall, featuring retail theft and pursuits that lead into the SEPTA transit system. Zone Five, which includes sports complexes, is characterized by long distances between stations. Zone Six includes the Temple University area, a scene of frequent crimes against students.

SEPTA officers must occasionally ask for assistance from their comrades in other zones. These calls are divided into two categories, "officer assists" and "officer backups." An "assist" requires officers to respond immediately. Often the only method available to get to the scene quickly is a run of five to eight city blocks. An officer responding to an "assist" must preserve enough energy to deal effectively with a situation once arriving on the scene. SEPTA averages about 380 running assists per year. "Backups" are not as critical as "assists," so officers generally use a "paced jog." SEPTA averages about 1,920 "backups" annually.

For help in attaining its fitness goals, SEPTA turned to Dr. Paul Davis, an acknowledged expert in the field who had recommended corrective measures for numerous law enforcement and government agencies. At the time Dr. Davis began his research for SEPTA, an officer's equipment load was 12 pounds; it is now nearly 26 pounds. Dr. Davis found that officers need "sound, intact, disease-free cardiovascular system[s]" to effectively perform their jobs. These requirements implicate aerobic capacity, *i.e.*, the ability of the body to utilize oxygen during sustained physical activities such as running, swimming, and cycling. Aerobic capacity is commonly measured in units of milliliters of oxygen per kilogram of body weight per minute — "mL/kg/min," or "mL."

SEPTA officers typically run or jog on a daily basis from three to eight city blocks for periods of three to ten minutes. They also engage in stair climbing, which requires a capacity of 54 mL. In light of this and other evidence, Dr. Davis concluded that SEPTA transit officers need an aerobic capacity of 50 mL. After determining that such a level would have a "draconian" effect on female applicants, however, Dr. Davis lowered his recommendation to 42.5 mL. That capacity could be demonstrated by running 1.5 miles in 12 minutes, a test that was adopted for applicants.

Dr. Davis had done a similar study for a fire department in St. Paul, Minnesota, which — in setting a standard of 45 mL — required applicants to run 1.5 miles in 11 minutes and 40 seconds. Eighty percent of male applicants and 76% of female applicants passed this test.

In addition to Dr. Davis' testimony, SEPTA also presented evidence from other experts to demonstrate a statistically significant correlation between aerobic capacity and the number of arrests made by individual SEPTA officers. Furthermore, of 207 commendations, 96% went to officers with an average capacity of 46 mL. Of these awards, 198 involved arrests, and 116 involved a foot pursuit, use of force or other physical exertion. Another study indicated that 51.9% of offense perpetrators had a capacity of 48 mL or higher, with only 27% having lower than a 42 mL rating.

The record demonstrates that a smaller percentage of female applicants passed the running test than males, but

that nearly all women who trained for it were able to pass. The named plaintiffs and some of the class members who failed demonstrated, for the most part, a "cavalier" attitude towards the running test. Videotapes showed some of these applicants walking at the halfway point, either because they were indifferent or unable to run for even that short a period of time. Thus, although there was a significant disparity between the pass-fail rates of male and female applicants, the extent of the difference appears to have been exaggerated to some extent by the approach taken by some of the applicants.

A physiologist, Dr. Lynda Ransdell, testified that 40% of all women starting at an aerobic capacity of 35 to 37 mL can train to pass the running test in eight weeks, and that 10% of all women between 20 and 29 years of age can do so without any training. She concluded that the average sedentary woman can achieve SEPTA's performance standard with only moderate training. SEPTA sent applicants a letter outlining recommended training techniques that Dr. Ransdell testified were adequate.

Plaintiffs introduced the testimony of Dr. William McArdle, who suggested the use of a "relative fitness" test in which all applicants would be required to meet the 50th percentile of aerobic capacity for their gender — approximately 42 mL for males, and 36 mL for females. However, Dr. Robert Moffatt, a defense expert who conducted tests of the aerobic capacity necessary to perform a SEPTA officer's duties, disagreed. He stated that female officers with a capacity of 36 mL would not be able to capably perform their duties after running to an "assist" or a "backup." Dr. Bernard Siskin, another defense expert, found that the arrest rate for females with a 36 mL capacity was significantly lower than that of males with a 42 mL capacity.

The District Court rejected Dr. McArdle's proposal because it would not serve SEPTA's business goal of providing a police force capable of performing the physical requirements of the job nearly as well as the existing test. Instead, the court found that "Dr. Davis' study, standing alone, met the professional standards for construct validation and satisfies defendant's burden of

demonstrating job relatedness and business necessity." Moreover, his study had sufficient empirical support for an aerobic capacity requirement of 42.5 mL.

II.

The dispute in this case centers on the applicable standard of business justification under the Civil Rights Act of 1991. *See* Pub. L. No. 102-166, Title I, § 105(a), 105 Stat. 1074-75 (adding 42 U.S.C. § 2000e-2(k)). The pertinent section provides: "An unlawful employment practice based on disparate impact is established . . . only if — [the] complaining party demonstrates that a respondent uses a particular employment practice that causes a disparate impact on the basis of . . . sex . . . and the respondent fails to demonstrate that the challenged practice is job related for the position in question and consistent with business necessity[.]" 42 U.S.C. § 2000e-2(k)(1)(A).

This addition to Title VII was passed in response to the Supreme Court's decision in *Wards Cove Packing Co., Inc. v. Atonio*, 490 U.S. 642 (1989). In that case, the Court held that after a plaintiff makes a prima facie showing of disparate impact, the defendant bears the burden to produce evidence of business justification. *See id.* at 659. The burden of persuasion, however, remains at all times with the plaintiff. *See id.* As to what showing would satisfy business justification, the Court held that "the dispositive issue is whether a challenged practice serves, in a significant way, the legitimate employment goals of the employer." *Id.* However, "there is no requirement that the challenged practice be 'essential' or 'indispensable' to the employer's business for it to pass muster." *Id.*

Some members of Congress were displeased with the result in *Wards Cove* and argued for a stricter standard of business justification based on their reading of pre-*Wards Cove* cases. After two years of legislative struggle, Congress and the President agreed upon a compromise bill. Whether the ambiguous language of the statute accomplished that purpose has been the subject of lively debate.[1]

---

1. In addition to the law review commentaries cited by the majority, see also Rosemary Alito, *Disparate Impact Discrimination Under the 1991*

The 1990 bill, which had been vetoed by the President,[2] had used the phrase "*required by* business necessity," rather than "*consistent with* business necessity," as used in the 1991 Act. The substitution of the word "consistent" was considered to indicate a standard less stringent than would "required." In that light, a fair reading of the 1991 Act is "the challenged practice is job related for the position in question and *in harmony* with business necessity."

It may fairly be said that the language ultimately adopted in the 1991 Act reflects an "agreement to disagree" and a return of the dispute to the courts for resolution. In short, unable to muster a veto-proof majority for either view, Congress "punted." This conclusion is underscored by Congress' highly unusual admonition that the courts consider only a designated "interpretive memorandum" as legislative history, rather than the more elaborate committee reports and other materials that customarily reveal the extent of the controversy between various views.

Civil Rights Act, 45 Rutgers L. Rev. 1011, 1033 (1993) ("Only . . . cases requiring proof of job-relatedness and a reasonable need for the challenged practice accord[] with both the statutory language of the 1991 Act and the applicable Supreme Court precedent."); Kingsley R. Browne, *The Civil Rights Act Of 1991: A "Quota Bill," A Codification Of Griggs, A Partial Return To Wards Cove, Or All Of The Above?*, 43 Case W. Res. L. Rev. 287, 349 (1993) ("business necessity" has the same meaning as the *Wards Cove* phrase "serves, in a significant way"); Linda Lye, Comment, *Title VII's Tangled Tale: The Erosion and Confusion of Disparate Impact and the Business Necessity Defense*, 19 Berkeley J. Employment & Lab. L. 315, 358 (1998) (a challenged practice must be a "reasonable predictor of effective performance of job duties," defined in light of "important business goals").

2. The fear of quota hiring was behind the President's refusal to sign earlier versions of the bill. *See Statement of President George Bush Upon Signing S. 1745*, *reprinted in* 1991 U.S.C.C.A.N. 768 (stating that the Act promotes the goals of ridding discrimination, allowing employers to hire on the "basis of merit and ability without the fear of unwarranted litigation," without leading to quotas or incentives for needless litigation). For a discussion of the drafting of the Civil Rights Act of 1991, see 2 Lex K. Larson, *Employment Discrimination* § 23.04[1] (2d ed. 1999). For analysis of the rejected 1990 bill, see Cynthia L. Alexander, *The Defeat of the Civil Rights Act of 1990: Wading Through the Rhetoric In Search of Compromise*, 44 Vand. L. Rev. 595 (1991).

See Pub. L. No. 102-166, Title I, § 105(b), 105 Stat. 1075. The interpretive memorandum states that: "The terms 'business necessity' and 'job related' are intended to reflect the concepts enunciated by the Supreme Court in *Griggs v. Duke Power Co.*, 401 U.S. 424 (1971), and in the other Supreme Court decisions prior to *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642 (1989)." 137 Cong. Rec. S15276 (daily ed. Oct. 25, 1991).

Congress' reference to the *Griggs* line of Supreme Court decisions, however, does little to clear the air because the language in those opinions has caused confusion.[3] The problem can ultimately be traced back to *Griggs* itself. In that case, which involved power-plant jobs, the Court held that a high school completion requirement and general intelligence tests that disproportionately disqualified black applicants were not significantly job related. The Court said: "The touchstone is business necessity." *Griggs*, 401 U.S. at 431. However, the very next sentence reads, "[i]f an employment practice . . . cannot be shown to be related to job performance, the practice is prohibited." *Id.* Thus, the Court speaks of both "necessity" and "job-relatedness" in the same breath.

In the following paragraph, we read that neither employment requirement is "shown to bear a demonstrable relationship to successful performance of the jobs for which it was used. Both were adopted . . . without meaningful study of their relationship to job-performance ability." *Id.* The Court also refers to "testing mechanisms [that are] unrelated to measuring job capability," "job-related tests," and states that "any given requirement must have a manifest relationship to the employment in question." *Id.* at 432-34, 436. Not once does the opinion repeat or expound upon "business necessity." Unquestionably, "job-relatedness" is *Griggs*' dominant thread.

The Court also cited with approval former EEOC Guideline 29 C.F.R. § 1607.4(c), which required employers

---

3. *See* Peter Brandon Bayer, *Mutable Characteristics and the Definition of Discrimination Under Title VII*, 20 U.C. Davis L. Rev. 769, 822 & n.213 (1987) ("Both the Supreme Court and lower court rulings offer a confusing patchwork of seemingly conflicting standards.").

to produce data "demonstrating that the test is predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated." *Id.* at 433 n.9.

The Court next visited the concept of business justification in *Albemarle Paper Co. v. Moody*, 422 U.S. 405 (1975), where a paper mill was using screening tests that had a disparate impact on black applicants. The issue, according to the Court, was whether the employer had shown the tests to be "job related." *Id.* at 408. The Court concluded that the employer's validation study was defective because it "involved no analysis of the attributes of, or the particular skills needed in, the studied job groups." *Id.* at 432. The Court was also critical of hiring decisions based on the subjective opinions of supervisors. *See id.* at 432-33.

The portion of *Albemarle* most relevant to the case at hand focused on whether tests that take into account capability for promotion may be utilized if such long-range requirements fulfill a "genuine business need." *Id.* at 434. The employer's validation study focused on the scores achieved by job groups near the top of the various lines of progression. The Court observed that those results did "not necessarily mean that the test, or some particular cutoff score on the test, is a permissible measure of the minimal qualifications of new workers entering lower level jobs." *Id.* at 434. Thus, the validation study was faulty because there had been "no clear showing that differential validation was not feasible for lower level jobs." *Id.* at 435.

The Court next considered appropriate criteria in *Washington v. Davis*, 426 U.S. 229 (1976), which involved written tests that allegedly had a discriminatory impact on black applicants for police officer positions. Although the suit was not brought under Title VII, the Court discussed *Griggs* and *Albemarle*. The district judge had concluded "that a positive relationship between the test and training-course performance was sufficient to validate the [test], wholly aside from its possible relationship to actual performance as a police officer." *Id.* at 250. Significantly, the Supreme Court remarked that such a conclusion was not foreclosed by either *Griggs* or *Albemarle* and "it seems

to us the much more sensible construction of the job-relatedness requirement." *Id.* at 250-51. Dismissing challenges to the test, the Court remarked that "some minimum verbal and communicative skill would be very useful, if not essential, to satisfactory progress in the training regimen." *Id.* at 250.

In another case, *Dothard v. Rawlinson*, 433 U.S. 321 (1977), the Court held that height and weight requirements for prison guards could not stand. The ruling was based on the employer's failure to produce any evidence to correlate those standards with "the requisite amount of strength thought essential to good job performance." *Id.* at 331. In a footnote, *Dothard* repeated *Griggs*' statement that "[t]he touchstone is business necessity," and further stated that "a discriminatory employment practice must be shown to be necessary to safe and efficient job performance to survive a Title VII challenge." *Id.* at 332 n.14. Earlier in the body of the opinion, the Court explained that the employer must show that a requirement has "'a manifest relationship to the employment in question.'" *Id.* at 329 (*quoting Griggs*, 401 U.S. at 432).

In yet another context, the Court upheld an employer's prohibition of employment to users of methadone, despite claims of disparate impact on blacks and Hispanics. *See New York City Transit Authority v. Beazer*, 440 U.S. 568, 587 (1979). To the Court, the employer's narcotics rule, even in its application to methadone users, was "job related." *Id.*

Beazer quoted the District Court's observation that "those goals [*i.e.*, safety and efficiency] are significantly served by—even if they do not require—[the employer's] rule as it applies to all methadone users including those who are seeking employment in non-safety-sensitive positions." *Id.* at 587 n.31. The Supreme Court concluded that "[t]he record thus demonstrates that [the employer's] rule bears a 'manifest relationship to the employment in question.'" *Id.* (*quoting Griggs*, 401 U.S. at 432).

The *Beazer* Court observed that most of the affected job positions were "attended by unusual hazards and must be performed by 'persons of maximum alertness and

competence.'" *Id.* at 571. Other positions were "critical" or "safety sensitive," and many involved "danger to [the employees] or to the public." *Id.*

III.

As the preceding sketch of pre-*Wards Cove* opinions demonstrates, the Supreme Court's articulations of the appropriate standards are far from clear. Phrases such as "business necessity," "demonstrable relationship to successful performance of the job," "manifest relationship to the employment in question," "genuine business needs," and "essential to good job performance," have been used interchangeably. These varying formulations bring to mind Justice Holmes' observation, "A word is not a crystal, transparent and unchanged, it is the skin of a living thought and may vary greatly in color and content according to the circumstances and the time in which it is used." *Towne v. Eisner*, 245 U.S. 418, 425 (1918).

My study of the standard for business justification as set forth by the Civil Rights Act of 1991 convinces me that it remains essentially the same as it was in the pre-*Wards Cove* era. However, other than its holding on burden of proof, it does not seem that *Wards Cove* was a revolutionary pronouncement. Until the Supreme Court reexamines the subject, however, courts will continue to struggle with the often inconsistent phraseology employed in *Griggs* and its progeny. The definition and application of the appropriate standard for business justification will depend on the context in which it is raised. There are significant factual differences in the cases that explain, to some extent, the differing formulations. *Albemarle* and *Griggs* applied greater scrutiny when the disparate impact affected entry to lower-level jobs, where it is fair to assume that no special qualifications would be generally expected.

In contrast, *Beazer* and *Washington* raised an additional important consideration — public safety. *Beazer* concerned jobs involving serious dangers to employees as well as to transit passengers. In *Washington*, a written test demonstrating an applicant's ability to complete police officer training was job-related, even apart from its

relationship to actual performance as a police officer. The impact of public safety concerns on employee qualifications is inescapable, and serves to differentiate those positions from lower-level, nonsafety-sensitive ones.[4]

The Courts of Appeals have explicitly recognized the relevance of safety considerations in a series of decisions beginning with *Spurlock v. United Airlines, Inc.*, 475 F.2d 216 (10th Cir. 1972). In that case, an airline required that applicants for flight officer positions have a college degree and a minimum of 500 flight hours. The Court, citing *Griggs*, held that where "the job clearly requires a high degree of skill and the economic and human risks involved in hiring an unqualified applicant are great, the employer bears a correspondingly lighter burden to show his employment criteria are job related." *Id.* at 219. Because, in the case of pilots, "[t]he risks involved in hiring an unqualified applicant are staggering . . . . [t]he courts . . . should proceed with great caution before requiring an employer to lower his pre-employment standards for such a job." *Id.*

Another leading case, *Davis v. City of Dallas*, 777 F.2d 205 (5th Cir. 1985), applied the *Spurlock* doctrine to criteria for hiring police officers. The City required a specific amount of college education, no history of recent marijuana usage, and a negative history of traffic violations. Despite findings of disparate impact, the Court upheld the requirements. Having reviewed the many cases following *Spurlock*, the Court had "no difficulty . . . equating the position of police officer in a major metropolitan area such as Dallas with other jobs that courts have found to involve the important public interest in safety." *Id.* at 215 (internal quotation marks omitted). The degree of public risk and responsibility alone "would warrant examination of the job relatedness of the . . . education requirement under the lighter standard imposed under *Spurlock* and its progeny." *Id.* at 215.

---

4. *See* Andrew C. Spiropoulos, *Defining the Business Necessity Defense to the Disparate Impact Cause of Action: Finding the Golden Mean*, 74 N.C. L. Rev. 1479 (1996).

Observing the nature of the positions at issue in *Griggs* and *Albemarle*, *Davis* noted that in neither case did the Supreme Court suggest that those jobs "were noteworthy for their dangerousness or importance to the public welfare." *Id.* at 210. In contrast, the employment under consideration in *Davis* directly implicated public safety concerns. *See id.* at 211. It is interesting that Justice Blackmun, in *Watson v. Fort Worth Bank & Trust*, 487 U.S. 977 (1988) (plurality op.), objecting to what he considered to be a tendency to weaken the employer's burden, cited *Davis* favorably, stating that "[t]he proper means of establishing business necessity will vary with the type and size of the business in question, as well as the particular job" in question. *Id.* at 1007. (Blackmun, J., concurring in part and concurring in the judgment).[5]

In a post-*Wards Cove* case involving firefighters, the Court of Appeals for the Eleventh Circuit noted that such "safety claims would afford the City an affirmative defense, for protecting employees from workplace hazards is a goal that, as a matter of law, has been found to qualify as an important business goal for Title VII purposes." *Fitzpatrick v. City of Atlanta*, 2 F.3d 1112, 1119 (11th Cir. 1993) (*citing Beazer*, 440 U.S. at 587 & n. 31; *Dothard*, 433 U.S. at 331 n. 14). Thus, "[m]easures demonstrably necessary to meeting the goal of ensuring worker safety are therefore deemed to be 'required by business necessity' under Title VII." *Id.*

In a similar case, the Court of Appeals for the Eighth Circuit wrote that "the law does not require the city to put the lives of [plaintiff] and his fellow firefighters at risk by taking the chance that he is fit for duty when solid scientific studies indicate that persons with test results similar to his are not." *Smith v. City of Des Moines*, 99 F.3d 1466, 1473 (8th Cir. 1996). Other Courts of Appeals have

---

5. In the analogous context of the defense of bona fide occupational qualification, the Supreme Court has stated: " 'The greater the safety factor, measured by the likelihood of harm and the probable severity of that harm in case of an accident, the more stringent may be the job qualifications . . . .' " *Western Air Lines, Inc. v. Criswell*, 472 U.S. 400, 413 (1985) (*quoting with approval Usery v. Tamiami Trail Tours, Inc.*, 531 F.2d 224, 236 (5th Cir. 1976)).

reached similar conclusions in cases involving safety-sensitive positions such as truck drivers, bus drivers, firefighters, and police officers.[6]

IV.

The issues that separate the parties are straightforward. Plaintiffs do not seriously contest the fact that aerobic capacity is a valid predictor of efficient job performance as a transit police officer. They do not challenge the finding that running for 1.5 miles is an effective way to measure aerobic capacity. Nor apparently do they suggest that 42.5 mL is an inappropriate cut-off for male applicants: they implicitly accept this standard by advancing Dr. McArdle's alternative test, which would use that score for males and a lower one for females.

Even the government plaintiff concedes that an employer may improve its workforce. U.S. Br. at 35 (*citing Griggs*, 401 U.S. at 431).[7] *Griggs*, in turn, stressed that tests "must

---

6. *See, e.g., York v. American Telephone & Telegraph Co.*, 95 F.3d 948, 952, 959 (10th Cir. 1996) (powerhouse operating engineers); *Zamlen v. City of Cleveland*, 906 F.2d 209, 217 (6th Cir. 1990) (firefighters); *Hamer v. City of Atlanta*, 872 F.2d 1521, 1535 (11th Cir. 1989) (firefighters); *Levin v. Delta Air Lines, Inc.*, 730 F.2d 994, 997-98 (5th Cir. 1984) (flight attendants); *Chrisner v. Complete Auto Transit, Inc.*, 645 F.2d 1251, 1261-63 (6th Cir. 1981) (truck yard employees); *Harriss v. Pan American World Airways, Inc.*, 649 F.2d 670, 676 (9th Cir. 1980) (flight attendants); *McCosh v. City of Grand Forks*, 628 F.2d 1058, 1063 (8th Cir. 1980) (police); *Boyd v. Ozark Air Lines, Inc.*, 568 F.2d 50, 54 (8th Cir. 1977) (airline pilots); *see also* Alito, *supra*, at 1033-35 & n.100.

7. It is interesting that in the legislative history of the original text of Title VII, congressional advocates argued that "title VII would not require, and no court could read title VII as requiring, an employer to lower or change the occupational qualifications he sets for his employees . . . ." 110 Cong. Rec. 7246-47 (April 8, 1964) (interpretive memorandum of Sen. Case). Senators Clark and Case stated that the "employer may set his qualifications as high as he likes . . . ." *Id.* at 7213 (April 8, 1964) (interpretive memorandum of Sens. Clark and Case). Senator Humphrey stated that "[t]he employer, not the Government, will establish the standards." *Id.* at 13088 (June 9, 1964). Thus, the legislative history of Title VII "clearly reveals that Congress was concerned about preserving employer freedom, and that it acted to mandate employer color-blindness with as little intrusion into the free enterprise system as possible." *Contreras v. City of Los Angeles*, 656 F.2d 1267, 1278 (9th Cir. 1981).

measure the person for the job and not the person in the abstract." *Griggs*, 401 U.S. at 436. SEPTA's running test attempts to do just that, *i.e.*, improve the caliber of its police force by selecting new hires to fit appropriately heightened performance standards.

A fair appraisal of the plaintiffs' objection is that the running test's cut-off requires female applicants to run faster than a majority of women can run without training. However, nearly all of the women who did train were able to pass the test. Also, not all males were able to pass, although their failure percentages were substantially lower.

Plaintiffs complain that SEPTA cannot point to any instances where a perpetrator of a crime got away, or an offense was committed because of an officer's lack of aerobic capacity. But as noted by *Fitzpatrick*, "[t]he mere absence of unfortunate incidents is not sufficient" to preclude a particular safety requirement because otherwise, such "measures could be instituted only once accidents had occurred rather than in order to avert accidents." *Fitzpatrick*, 2 F.3d at 1120-21.

Here, where applicants have it within their power to prepare for the running test, they may properly be expected to do so. In view of the important public safety concerns at issue, it is not unreasonable to expect all applicants — female or male — to take the necessary steps in order to qualify for the positions.

The District Court's conclusions must be appraised against this background. The trial was lengthy and the evidence extensive. The findings of fact on job needs with respect to aerobic capacity are not clearly erroneous. This conclusion is mandated by the standard that clear error exists only when, on the entire evidence, a court is left with the definite, firm conviction that a mistake has been committed. *See Anderson v. City of Bessemer City*, 470 U.S. 564, 573 (1985). If the account of the District Court is "plausible in light of the record viewed in its entirety," we may not reverse even if we are convinced that had we "been sitting as the trier of fact, [we] would have weighed the evidence differently." *Id.* at 574.

Moreover, "[w]here there are two permissible views of the evidence, the factfinder's choice between them cannot be clearly erroneous." *Id.* "This is so even when the district court's findings . . . are based instead on physical or documentary evidence or inferences from other facts." *Id.* Where findings are based on credibility determinations, appellate review accords even greater deference to the findings of the District Court. *See id.* at 575. Courts routinely hold that business justification is reviewed for clear error. *See, e.g., Davis*, 777 F.2d at 208 & n.1; *Spurlock*, 475 F.2d at 219-20. I accept, therefore, that 42.5 mL is an appropriate level for the position of a SEPTA officer, that it is reasonable, and that it is attainable by otherwise physically fit female applicants with moderate training.

The question then, is whether SEPTA's standard is permissible under the terms of the Civil Rights Act of 1991 and the relevant precedents. The District Court rejected the plaintiffs' contention that "business necessity" under the statute is governed by a footnote in *Dothard* that states: "[A] discriminatory employment practice must be shown to be necessary to safe and efficient job performance . . . ." *Dothard*, 433 U.S. at 332 n.14. Rather, looking to *Griggs* and *Beazer*, the District Court stated that SEPTA need only show that its tests "significantly serve, but are neither required by nor necessary to, the employer's legitimate business interests" — in other words, that it "bears a manifest relationship" to the employment in question.

In disagreeing with the criteria used by the District Court, the majority holds that "a discriminatory cutoff score is impermissible unless shown to measure the minimum qualifications necessary for successful performance of the job in question." The difficulties presented by this standard are illustrated by the testimony of Dr. McArdle, the plaintiffs' expert. In essence, he proposed that female applicants be expected to meet 50% of their aerobic capacity, translating to 36 mL, but that males continue at the 50% level of 42.5 mL. That standard would, of course, have less adverse impact on women, but according to the findings of the District Court, would also have a detrimental impact on the effectiveness of the SEPTA transit police.

With this in mind, I cannot agree that the majority's standard is the correct one for this case. Reducing standards towards the lowest common denominator is particularly inappropriate for a police force. Undoubtedly, candidates who fail the running test — female or male — may have other qualities of particular value to SEPTA, but they must possess the requisite aerobic capacity as well. No matter how laudable it is to reduce job discrimination, to achieve this goal by lowering important public safety standards presents an unacceptable risk.

Aerobic capacity is an objective, measurable factor which gauges the ability of a human being to perform physical activity. The aerobic demands on the human system are affected by absolutes such as the distance traveled, the speed, the number of steps to be climbed, and similar factors. Governmental agency pronouncements will not shorten distances, reduce the number of steps, or decrease the aerobic capacity of perpetrators to match the reduced standards of officers, male or female.[8] Some males and more females cannot meet the necessary requirements. Based on the facts established at trial, those individuals simply cannot perform the job efficiently. To the extent that they cannot, their hire adversely affects public safety.

---

8. Although the government is a plaintiff in this dispute, I would note that some agencies take a somewhat different tack on the issue of aerobic fitness. The U.S. Forest Service, for instance, requires firefighters to have an aerobic capacity of 45 to 48 mL, and recommends one of up to 50. *See* United States Department of Agriculture, Forest Service, Technology & Development Program, *Fitness and Work Capacity* 51 (2d ed. 1997). Notably, that agency currently uses a 1.5 mile run test. *See id.* at 50-51.

Also, the Presidential Physical Fitness Award is available to children who meet the 85th percentile of fitness by meeting target levels in events such as a one-mile run. *See* Qualifying Standards (updated Oct. 15, 1998). <http://www.indiana.edu/~preschal/qualifying.html>.

The Centers For Disease Control and Prevention lament that more than 60% of U.S. adults do not engage in the recommended amount of activity, and 25% are not active at all. *See* Physical Activity and Health, Adults (viewed May 7, 1999) <http://www.cdc.gov/nccdphp/sgr/adults.htm>.

The current Uniform Guidelines on Employee Selection Procedures, 29 C.F.R. § 1607 ("EEOC Guidelines"), are not as strict as the standard suggested by the majority. In discussing cut-off scores, the Guidelines explicitly state that "they should normally be set so as to be reasonable and consistent with normal expectations of acceptable proficiency within the work force." 29 C.F.R. § 1607.5(H) (1998). Further, the EEOC Guidelines standard — "predictive of or significantly correlated with important elements" — has been cited by the Supreme Court with approval on several occasions. *See Albemarle*, 422 U.S. at 431 (*quoting* former 29 C.F.R. § 1607.4(c)); *Griggs*, 401 U.S. at 433 n.9 (*quoting* same); *see also* 29 C.F.R. § 1607.5(B) (1998).

Further, *Albemarle*'s reference to "minimal qualifications" was directed only to the inappropriateness of using a test geared towards higher-level jobs as a screen for entry-level positions. *See Albemarle*, 422 U.S. at 434. This holding, which is minimally relevant to the matter at hand, is doubly inapplicable when the job affects public safety. *See Davis*, 777 F.2d at 211 n.5.[9]

I see no need to remand this case to the District Court. Whatever standard is used, the findings of fact require an affirmance. Although the District Court rejected the plaintiffs' argument that the *Dothard* footnote, rather than *Beazer*, supplied the proper standard, the factual findings make it clear that under either formulation, the District Court reached the correct result.

The *Dothard* footnote states that the challenged practice must be "necessary to safe and efficient job performance." *Dothard*, 433 U.S. at 331 n.14. The District Court, also in a footnote, wrote "physical fitness is only one trait or ability required of SEPTA officers, [but] it is a trait or ability that is necessary for and critical to the successful performance of the job, and thus SEPTA should be able to test for such

---

9. The plaintiffs also suggest that SEPTA's validation studies were insufficient. However, strict compliance with the EEOC Guidelines is not necessary in all cases. *See Beazer*, 440 U.S. at 587 n.31; *Washington*, 426 U.S. at 250-51. In cases involving public safety, courts have held that empirical validation is not required. *See Boyd*, 568 F.2d at 54.

a trait." This finding more than complies with *Dothard*'s footnote by concluding that not only is physical fitness "necessary" to safe and efficient job performance as SEPTA officers, but that it is "critical" to successful performance of these jobs. Moreover, the finding clearly meets even the criterion that cut-off scores "measure the minimum qualifications necessary for *successful performance* of the job." (emphasis added).

Nor can there be any doubt that the factual findings here satisfy *Griggs*' requirement of "business necessity." Unquestionably, SEPTA's test is job-related and there can be no doubt that physical fitness, and particularly aerobic capacity, is necessary for adequate performance of the job of a SEPTA transit officer. The findings are convincing that 42.5 mL is a reasonable cut-off point for determining the physical ability necessary for successful performance of the job. Consequently, even under the plaintiffs' reading of the 1991 Act, which relies so much on *Dothard*, the judgment in favor of the defendant should be affirmed.

To my mind, the correct standard for this case is that of *Spurlock-Davis*, one that places greater emphasis on the safety of the public and fellow officers. I have no doubt that this line of cases survives the Civil Rights Act of 1991, because those opinions — as noted in Congress' "interpretive memorandum" — "reflect the concepts enunciated" in Supreme Court decisions prior to *Wards Cove*. *See Watson*, 487 U.S. at 998; *Beazer*, 440 U.S. at 587 n.31; *Washington*, 426 U.S. at 250; *Smith*, 99 F.3d at 1473; *Fitzpatrick*, 2 F.3d at 1119. Safety concerns are clearly "concepts" considered by the Supreme Court and applied in various factual circumstances by the Courts of Appeals, both in pre- and post-*Wards Cove* cases. Nothing in the legislative history casts any doubt on the continued viability of these opinions.

Although it did not cite *Spurlock-Davis*, the District Court stated in its conclusions of law that "employers such as SEPTA should be encouraged to improve the efficiency of its workforce, especially where public safety is implicated by the particular job as it is with SEPTA." More emphatically, it stated that "[t]he Court simply will not condone dilution of readily obtainable physical abilities standards that serve

to protect the public safety in order to allow unfit candidates, whether they are male or female, to become SEPTA transit police officers."

Although the District Court only inferentially applied *Spurlock-Davis*, I would do so explicitly and affirm the judgment on that basis.[10] Here, the record supplies ample evidence about safety concerns related to the performance of SEPTA officers. In cases such as these, courts should decline to lower standards in an effort to reduce disparate impact when that goal comes at the expense of public safety. Due deference should be afforded to the experience of specialized employers in setting appropriate requirements for safety-sensitive positions.

V.

The Lanning appellants propose a number of alternative practices that they suggest would have a lesser disparate impact while still serving SEPTA's goals. First, they suggest that SEPTA select medically fit applicants who pass fitness requirements at the end of their training at the Philadelphia Police Academy. Second, as noted earlier, they argue in favor of a relative fitness test (*i.e.*, one with a lower cut-off point for females). Third, they prompt SEPTA to propose an alternative.

For plaintiffs to establish a satisfactory alternative, they must "make[ ] the demonstration described in [42 U.S.C. § 2000e-2(k)(1)(C)] with respect to an alternative employment practice and [establish that] the [employer] refuses to adopt such alternative employment practice." 42 U.S.C. § 2000e-2(k)(1)(A)(ii). To meet this burden, the plaintiffs' proposed alternatives must have less disparate impact and "also serve the employer's legitimate interest in 'efficient and trustworthy workmanship.' " *Albemarle*, 422 U.S. at 425; *see also NAACP v. Medical Ctr., Inc.*, 657 F.2d 1322, 1336 n.17 (3d Cir. 1981) (en banc). As stated in

---

10. An order of the District Court may be affirmed on alternative grounds where the judgment is supported by the record below. *See Guthrie v. Lady Jane Collieries, Inc.*, 722 F.2d 1141, 1144-45 & n.1 (3d Cir. 1983) (citing *Helvering v. Gowran*, 302 U.S. 238, 245 (1937)).

Watson, the alternative test must "be equally as effective as the challenged practice in serving the employer's legitimate business goals." *Watson*, 487 U.S. at 998.

The District Court found that none of the plaintiffs' proposals served SEPTA's legitimate interest in having a more physically fit work force. If SEPTA may require an aerobic capacity of 42.5 mL *after* training at the police academy, as plaintiffs propose, it is unclear how that practice would be any less discriminatory than requiring it *before* hire. In short, that plan would simply require that training be on "company time" rather than on that of the applicants.

As to the relative fitness test proposed by the plaintiffs' expert, the factual findings demonstrate that officers with a capacity of 36 mL do not serve SEPTA's needs as well as the required standard of 42.5 mL.[11] Finally, the proposal that SEPTA come forward with an alternative is not an alternative at all. Thus, plaintiffs have failed to meet their burden to establish an alternative employment practice.

I would affirm the judgment of the District Court.

A True Copy:
Teste:

Clerk of the United States Court of Appeals
for the Third Circuit

---

11. The Civil Rights Act of 1991 presents another potential barrier to the relative fitness test. Subsection 2000e-2(l) prohibits "in connection with the selection or referral of applicants or candidates for employment . . . to . . . use different cutoff scores for . . . employment related tests on the basis of . . . sex[.]" By its plain language, 42 U.S.C. § 2000e-2(l) arguably prohibits a relative fitness test. The District Court concluded that this provision did not apply. I have some doubt on that ruling, but need not reach that issue because I would affirm on other bases.

APPENDIX B

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF PENNSYLVANIA

CATHERINE NATSU LANNING, ET AL., *Plaintiffs,*	:	Civil Action
v.	:	
SOUTHEASTERN PENNSYLVANIA TRANSPORTATION AUTHORITY, *Defendant.*	:	No. 97-0593

UNITED STATES OF AMERICA, *Plaintiff,*	:	Civil Action
v.	:	
SOUTHEASTERN PENNSYLVANIA TRANSPORTATION AUTHORITY, *Defendant.*	:	No. 97-1161

Newcomer, J. June __, 1998

FINDINGS OF FACT AND CONCLUSIONS OF LAW

After a bench trial of this case on January 7-9, 12-16, 20, 21, 23, and 26, 1998, and after considering the testimony of the witnesses, the admitted exhibits and the arguments of counsel, and the parties' post-trial submissions, the Court makes the following findings of fact and conclusions of law.

FINDINGS OF FACT

A. The Parties

1. The individual named plaintiffs in Civil Action No. 97-0593 are Catherine Natsu Lanning, Denise Dougherty, Altovise Love, Belinda Kelly Dodson and Lynne Zirilli (formerly Lynne Carapucci).

2. The plaintiff class in Civil Action No. 97-0593 was certified by the Court on August 10, 1997 pursuant to Fed. R. Civ. P. 23(b)(2) and is defined as:

all 1993 female applicants, 1996 female applicants and future female applicants for employment as SEPTA police officers who have been or will be denied employment by reason of their inability to meet the physical entrance requirement of running 1.5 miles in 12 minutes or less.

3. The plaintiff in Civil Action No. 97-1161 is the United States of America ("United States").

4. Defendant is the Southeastern Pennsylvania Transportation Authority ("SEPTA"), a regional mass transit authority that currently operates under authority of the Metropolitan Transportation Authorities Act, 74 Pa. Con. Stat. Ann. §§ 1701, *et seq.*, which confers upon SEPTA "the public powers of the Commonwealth as an agency and instrumentality thereof." *Id.* § 1711(a). SEPTA's principal office is located at 1234 Market Street, Philadelphia, Pennsylvania 19107. Defendant SEPTA is a person within the meaning of 42 U.S.C. § 2000e(a) and an employer within the meaning of 42 U.S.C. § 2000e(b).

5. This Court has jurisdiction of Civil Action No. 97-0593 under 42 U.S.C. §§ 2000e-1, *et seq.* and 28 U.S.C. §§ 1331, 1343 (a)(1), (a)(3), (a)(4).

6. This Court has jurisdiction of Civil Action No. 97-1161 under 42 U.S.C. § 2000e-6(b), 28 U.S.C. §§ 1343(a)(3), 1345.

7. Venue is proper in Civil Action Nos. 97-0593 and 97-1161 under 28 U.S.C. § 1391(b).

B. Procedural History

8. In accordance with Section 707 of Title VII of the Civil Rights Act of 1964, as amended, 42 U.S.C. § 2000e, *et seq.* ("Title VII"), the United States, through the Department of Justice, provided written notice to defendant SEPTA in March 1996, that it was conducting an investigation of SEPTA's employment practices.

9. Prior to the filing of the United States' Complaint in Civil Action No. 97-1161, the Attorney General of the United States found reasonable cause to believe that SEPTA was engaged in a pattern or practice of employment discrimination

against women through the use of its physical fitness test given to transit police officer applicants in violation of Section 707 of Title VII.

10. Prior to the filing of the United States' Complaint, the United States provided written notice, in February 1997, to defendant SEPTA of the Attorney General's reasonable cause determination and thereafter unsuccessfully attempted to resolve this matter through negotiation prior to filing suit.

11. All conditions precedent to the filing of suit in Civil Action No. 97-1161 have been met.

12. In 1993, each of the individual named plaintiffs applied for a position as a SEPTA transit police officer. In October 1993, each of the individual named plaintiffs took a written examination administered by SEPTA for the position of transit police officer, and each individual was subsequently notified by SEPTA of her eligibility to proceed to the next phase of the selection process.

13. On October 30, 1993, each of the individual named plaintiffs participated in a 1.5 mile run as part of the physical entrance test. None of the individual named plaintiffs completed the 1.5 mile run within the required 12 minute cutoff time set by SEPTA.

14. Each of the individual named plaintiffs was subsequently informed that because of her failure to complete the 1.5 mile run in 12 minutes or less, she was being rejected by SEPTA and would not be permitted to continue in the selection process.

15. In April 1994, each of the individual named plaintiffs filed an administrative charge of discrimination with the Pennsylvania Human Relations Commission ("PHRC") and the U.S. Equal Employment Opportunity Commission ("EEOC") challenging SEPTA's physical fitness test as discriminatory against women. On February 1, 1996, the PHRC issued a finding of probable cause that discrimination occurred with respect to each of the five individual charges of discrimination. The parties engaged in conciliation but were unable to resolve the matter prior to the filing of Civil Action No. 97-0593. On December 11, 1996, the EEOC issued a notice of right to sue to each of the five individual plaintiffs.

16. All administrative prerequisites to filing suit in Civil Action No. 97-0593 by the *Lanning* plaintiffs have been satisfied.

17. On January 25, 1997, the individual plaintiffs filed their class action Complaint in Civil Action No. 97-0593 against SEPTA and SEPTA Police Chief Richard J. Evans alleging that SEPTA was engaged in a pattern or practice of discrimination against female transit police officer candidates by using a physical fitness test that disproportionately excludes women and was neither predictive of successful job performance nor consistent with SEPTA's legitimate business necessity. The *Lanning* plaintiffs further alleged that there existed less discriminatory alternative selection devices that would serve SEPTA's legitimate business interest but that would have less or no adverse impact against women. The Complaint asserted causes of action under Title VII, the Pennsylvania Human Relations Act ("PHRA"), and 42 U.S.C. § 1983.

18. On February 18, 1997, the United States filed its Complaint in Civil Action No. 97-1161 pursuant to Section 707 of Title VII, 42 U.S.C. § 2000e-6, alleging that SEPTA was engaged in a pattern or practice of discrimination against female transit police officer candidates by using a physical fitness test, including but not limited to a 1.5 mile run, that disproportionately excludes women and was neither predictive of successful job performance nor consistent with SEPTA's legitimate business necessity. The United States further alleged that less discriminatory alternative selection devices existed that would serve SEPTA's legitimate business interest but that would have less or no adverse impact against women.

19. On April 21, 1997, the Court consolidated Civil Action Nos. 97-0593 and 97-1161 for all purposes up to and including trial.

20. On August 10, 1997, the Court dismissed the *Lanning* plaintiffs' claims under 42 U.S.C. § 1983.

21. On November 25, 1997, the Court granted the motion of the *Lanning* plaintiffs to amend their complaint to withdraw their claims of intentional discrimination under Title VII and the PHRA as well as their claim of disparate impact under the PHRA.

C. SEPTA's Selection Procedure for Transit Officers

1. The Impetus for SEPTA's Physical Fitness Test

22. In January 1989, Howard Roberts was hired by SEPTA as the Deputy General Manager. As the Deputy General Manager, Mr. Roberts was entrusted with managing the SEPTA Transit Police Department.

23. Shortly after his arrival in 1989, Mr. Roberts became aware of significant problems with the SEPTA Transit Police Department. Most notably, Mr. Roberts noticed that the SEPTA Transit Police Department was unable to control crime on SEPTA property and that problems existed with the physical fitness and capabilities training of its transit police officers.

24. At the time Mr. Roberts arrived at SEPTA, there were no physical fitness standards or physical training programs in place for SEPTA officers. As a result, there were instances where officers were injured, and there were numerous cases of police brutality that were caused by officers retaliating against persons who had previously assaulted physically unfit police officers.

25. Mr. Roberts noted that "crime statistics were very, very bad, officers for the most part arrived at crimes after they had taken place and basically did reports and turned them in." In essence, the SEPTA Transit Police Department was not preventing crime, rather it was merely reporting crime that occurred on SEPTA property.

26. In response to these problems, SEPTA initiated a complete overhaul of the police department under the direction of Mr. Roberts; its goal was to make the subways on the SEPTA system the "safest place in the city." This overhaul included the announcement that transit police were to be primarily dedicated to the subway and were not to serve as guards to protect personal or physical property at depots. SEPTA increased the number of officers from 96 to nearly 200 and introduced a "zone concept" for the area they patrolled.

a. The zone concept was implemented to decentralize the officers and place them out in the communities where they patrol. At any given time throughout the course of

the year, SEPTA police officers are assigned to a particular zone — in total, SEPTA has eight zones.

b. In a typical zone, there is a Lieutenant that commands the zone and two Sergeants. There are three tours everyday, *i.e.*, three shifts in a 24-hour period.

c. A beat is an assigned patrol area within the zone. The beats are reassigned on a daily or weekly basis to familiarize the officer with the entire zone of responsibility and to permit the officer to establish a relationship with the various cashiers and passengers throughout the zone.

d. The officers are deployed alone and on foot. When manpower permits, the beats are assigned in overlapping fashion to minimize the distances that officers will have to run to effectuate "officer backups" and "officer assists." Absent full availability of all zone officers, officer backups or officer assists routinely come from two stations away. There is usually one vehicle patrolling in each zone. However, due to the age of the vehicles and due to other uses of the vehicle, such as the transporting of prisoners, foot patrol officers cannot rely on backup coming from the patrol vehicle.

e. Upon arriving at the first station at the beginning of their shift, officers inspect for hazardous conditions, observe the station, employees and passengers, inspect the cashier areas and proceed to the next station to repeat the process.

f. The terrain of SEPTA's system varies greatly throughout the eight zones. Much of the terrain adds additional physical demands on the foot patrol officer. For example, Zone 1 is predominantly outdoors, exposing an officer to the elements for the entire eight hour tour. Zone 1 features the Market-Frankford elevated line which necessitates a climb of 30 to 50 steps from street level to the platform area. There is also a catwalk that officers sometimes use to run from station to station. The SEPTA officers also work in an environment that often causes them to effectuate their duties in crowds and in close quarters. This presents a heightened danger to the solo

patrol officer because crowds in the vicinity of an arrest or pedestrian stop will often side with the perpetrator over the officer. For example, Lt. Timothy Maslin has been struck from behind in a crowd situation. Moreover, after a SEPTA officer effectuates an arrest, he or she, unlike a car-based patrol officer, remains immersed in the same system on foot, exposing himself or herself to hostile crowds.

g. Zone 2 is an underground zone. Zone 3 is a mixture of above and below ground locations. Zone 3 also borders a large shopping mall, and therefore features more retail theft and pursuits that lead into the SEPTA system.

h. Zone 4 runs from Huntington Station to Bridge-Pratt on the Market-Frankford Line. It is an elevated portion of the system. This zone shuts down its stations at 8:00 p.m. and places riders on shuttle buses that are staffed with a special bus detail unit to protect the riders.

i. Zone 5 features large distances between stations, requiring officers to run longer for foot-based officer backups and officer assists. Zone 5 also features the Philadelphia sports complexes — Veterans Stadium, the CoreStates Center and the Spectrum.

j. Zone 6 is similar to the other zones except for the Ridge Spur that runs from Chinatown to Erie Station. The biggest area of concern is the Temple University area and the prevention of crime against students.

k. Zone 7 runs from Allegheny Station to Fern Rock Transportation Center on the Broad Street Line and features a high crime area and two major high schools. SEPTA emphasizes prevention of vandalism, graffiti and theft from the Fern Rock train yard. Lt. Maslin emphasized that due to the size of the train yards, an officer is not always able to pinpoint his or her location in the yard when stopping a trespasser.

l. Zone 8 deals strictly with Regional Rail operations covering 2,200 square miles and five counties. The officers in Zone 8 are often outside of SEPTA's radio frequency;

therefore, the officers patrol in pairs on occasion. Market East Station, Suburban Square and University City Station are also in Zone 8.

m. During their tours, SEPTA officers frequently respond to officer assist or officer backup calls. An officer assist call requires other officers to respond immediately to another officer's call for assistance — the responding officers are expected to use any means to get to the officer requiring assistance. An officer backup call also requires other officers to respond to the officer requesting assistance; however, the officers responding to a backup call do not have to arrive as quickly as they would for an officer assist situation. In essence, an officer assist call indicates that an officer is involved in or about to become involved in a potentially hostile or life- or property-threatening situation.

n. SEPTA officers have only two means by which to respond to officer backup and officer assist calls: (1) ride a train to the location where help is needed, if a train is available; or (2) run to the location where assistance is needed. Backups are run as paced jogs. Assists are paced runs with the goal of maintaining enough reserve energy to engage in any necessary struggling at the location of the call. SEPTA averages about 4 running assist responses per zone per month. Over eight zones, this is approximately 32 running assists per month or approximately 380 running assists per year. SEPTA averages about 20 running backups per zone per month. Over eight zones, this is approximately 160 running backups per month or approximately 1,920 running backups per year.

o. SEPTA backup calls are not broadcast city-wide. SEPTA does not rely on any other police department or jurisdiction to provide backup to its officers. Assist calls are relayed from SEPTA's dispatcher to "J Band," a city-wide frequency that is used to seek assistance from any available jurisdiction. In some cases, police officers from other jurisdictions, most notably the Philadelphia Police

Department, will arrive at the scene of a SEPTA officer backup or assist call. SEPTA officers, conversely, respond to officer assist calls from all other jurisdictions.

p. Notably, SEPTA officers are not always able to receive backup because certain underground locations are "dead zones" of steel and concrete that block the radio transmission frequencies from escaping out to the dispatcher's office.

27. While increasing the number of police officers from 96 to 200, SEPTA noticed that a large number of applicants were retirees from the Philadelphia Police Department. Because SEPTA was concerned about the physical fitness of these retirees, and because of the poor physical fitness of its incumbent force, SEPTA imposed the requirement that applicants to the position of transit police officer be 35 years of age or younger.

28. In response to a claim of age discrimination, and in accordance with a recommendation from the EEOC, SEPTA abandoned its age-based hiring and instead decided to commission a study to develop job-related physical fitness tests to be used for the testing of applicants for the transit police officer position.

2. Dr. Davis' Development of a Physical Fitness Test

29. In 1991, SEPTA hired Dr. Paul Davis to develop and validate a physical fitness test. Dr. Davis is a preeminent expert in the field of physical fitness and employment testing, and he has designed numerous fitness tests for law enforcement agencies, fire departments, armed services personnel and other entities engaged in the protection of the public.

30. In developing physical abilities testing, Dr. Davis uses a "research design approach," and applies criterion-related, construct and content validation strategies.[1] Dr. Davis believes

---

1. Courts and the psychological profession generally recognize three validation studies: content validity, criterion-related validity and construct validity. *Washington v. Davis*, 426 U.S. 229, 96 S. Ct. 2040, 48 L. Ed. 2d 597 (1976). *See also Uniform Guidelines on Employee Selection Procedures*, 29 C.F.R. § 1607, *et seq.*, (*"Uniform Guidelines"*). In general, test validation is the process by which it is determined whether the inferences that the employer draws from results on a selection device are appropriate and

that the rationale for physical abilities testing[2] is to ensure that there is an appropriate match between the requirements of the job and the individual who is applying for that position.

31. Prior to SEPTA, Dr. Davis had experience with developing physical abilities tests for numerous police and fire departments, approximately 70 different organizations.

32. Dr. Davis has also participated in a project for the United States Marine Corps, spending six years attending formal military schools for desert, mountain and jungle warfare and amphibious operations for the purpose of developing physical fitness tests for the Marine Corps.

33. Some job task analyses[3] that Dr. Davis has performed have included ride-along programs with the Indiana State Police and with fish and game officers in Wyoming. Job task analysis can also include observation of and interviews with employees.

34. Prior to SEPTA, Dr. Davis also completed a police project for Anne Arundel County, Maryland, as well as a project for the Department of Public Safety of Oakwood, Ohio, which included firefighters and police officers.

35. In Oakwood, Dr. Davis spent the better part of a week on a ride-along with the police officers, observing the kinds of activities that the officers engage in throughout the course of a day.

36. The Anne Arundel County Project was for the Department of Corrections, the fire department and the police

---

NOTES *(Continued)*

meaningful. That is, test validation attempts to determine whether (and the degree to which) persons who are selected by a test will be successful performers on the job, and whether those who are not selected would not have been successful performers on the job.

2. Throughout these Findings of Fact and the Conclusions of Law, the Court will use the terms "physical abilities test (or testing)" and "physical fitness test (or testing)" interchangeably.

3. All test validation begins with a job analysis in which an effort is made to determine the specific knowledge, skills and abilities which are important to successful performance of the position in question. *Uniform Guidelines* § 14A.

department. Dr. Davis developed a physical abilities test for all three units, and he performed a job task analysis for all three units.

37. With regard to SEPTA, Dr. Davis was contacted by Dr. Louis Vanderbeek, the Director of Medical Programs for SEPTA, to develop a physical fitness program for SEPTA. Early in the project, Dr. Davis met with Judith Pierce, the Assistant General Manager of SEPTA, Ronald Sharpe, the Chief of the SEPTA Transit Police Department, and other SEPTA officials to understand exactly what SEPTA's objective was with respect to developing a physical fitness test.

38. Based upon his meetings with SEPTA officials, Dr. Davis came to understand that SEPTA was trying to enhance the level of fitness, physical vigor and general productivity of its police force; SEPTA also wanted medical criteria from which it could make informed decisions regarding such issues as return to duty, hiring and retirement. From these interviews, Dr. Davis also discovered that crime was rampant on the SEPTA system and that there were questions about safety for the ridership of SEPTA. Davis further learned that SEPTA wanted to remedy this situation and that SEPTA believed that improving the physical fitness of its police force was one of the best methods to achieve such a goal.

39. In addition to these meetings, Dr. Davis went on a ride-along with SEPTA transit police officers and went out on the trains, covering virtually all of SEPTA's properties and obtaining a strong perspective of the expectations for transit officers. Dr. Davis spent approximately 20 hours traveling within the transit system over the course of approximately two days.

40. In these 20 hours, Dr. Davis learned that SEPTA had a foot-based patrol and that there was a probation against sitting down so that the officers will always be in a state of readiness. Dr. Davis also discovered that SEPTA radio communications were at times unreliable because many of the transmissions were underground, causing a lack of instant contact for backup in many cases. Dr. Davis further noticed the distances that officers have to cover by foot and the number of stairs that officers have to surmount on a daily basis.

41. Davis learned that a zone system had been established and that officers were expected to patrol aggressively within their assigned beat. The foot patrol officers are expected to patrol a two to three station area. If an officer assist call was received and no train was available, the SEPTA officer would have to "hit the bricks" and go to the next station which can be anywhere from five to eight blocks away.

42. Dr. Davis further discovered that SEPTA officers encounter the equivalent of a five story walk-up or walk-down of stairs on a daily basis.

43. What distinguishes the essential tasks or functions required of a SEPTA transit officer from the essential tasks required of police officers from other law enforcement agencies is that all of the activities take place on foot; therefore, the expectation is that SEPTA officers will have to move, run and walk with a higher degree of frequency on a daily basis more than other law enforcement officers. Dr. Davis found that a SEPTA officer would need a "sound, intact, disease-free cardiovascular system" to effectively perform their job. Dr. Davis testified that having such a cardiovascular system translates into aerobic capacity.

44. Aerobic capacity is the ability of the body to utilize oxygen during sustained physical activities such as running, swimming and cycling. Aerobic capacity is commonly measured in units of milliliters of oxygen per kilogram of body weight per minute ("mL/kg/min"). Aerobic capacity is often referred to as "VO2 max" (maximum value of oxygen). The more milliliters of oxygen per kilogram of body weight a person is capable of consuming during high-intensity sustained physical effort, the higher his or her VO2 max score.

45. In an effort to determine what physical abilities are required to perform as a SEPTA officer, Dr. Davis conducted a job task analysis specifically for SEPTA, using a Delphi session.

46. A Delphi session is one way to do a job task analysis. Instead of sending voluminous surveys to be administered to the SEPTA officers, Dr. Davis ascertained more quickly and more efficiently the same information employing the Delphi technique.

47. Through the Delphi technique, officers arrive at a consensus opinion about some issue (e.g., what tasks do officers encounter on a daily basis) that would be fairly close to the truth on the basis of everyone's experience on the job as a SEPTA officer. The officers that participated in the Davis' Delphi session, subject matter experts ("SMEs"), had a cumulative experience of over 100 police years that proved to be invaluable to Dr. Davis.

48. Within the context of law enforcement, a subject matter expert ("SME") is an individual who has considerable experience within that profession, having attended the requisite schools, perhaps advanced schools, and training programs, as well as experience on the job.

49. Twenty SMEs from SEPTA participated in the Delphi session conducted by Dr. Davis. Dr. Davis requested that the SMEs at SEPTA have at least five years of experience on the job. The SMEs were also representative of the demographics of SEPTA with regard to gender and race.

50. The Delphi session that was undertaken at SEPTA began by presenting the SMEs with a list of physical tasks that police officers engage in while performing their jobs. This list was called the Taxonomy of Physical Tasks ("Taxonomy"). Dr. Davis prepared this list by perusing descriptions of law enforcement activities and drawing upon his own personal experience observing law enforcement agencies. The assumption underlying the list is that if you can do the hardest tasks on the list, then you can do all the other tasks that are listed below that task on the list.

51. In his study, Dr. Davis presented the Taxonomy to the SMEs and asked them to indicate whether they ever engaged in the tasks while performing their duties with SEPTA. After the SMEs indicated whether they engaged in the activities, Dr. Davis verified that the list was reasonably comprehensive.

52. The SMEs then determined the relative importance of the tasks. Dr. Davis presented the SMEs with a scale that ranked the criticality of the particular physical task from one to five or six - one being the least critical and five or six being the most critical. Dr. Davis was not sure whether the top of the scale was five or six because he did not retain the scales

used in his study. Dr. Davis, however, did indicate that this discrepancy would not upset his results because a ranking of either five or six would mean that the SMEs viewed the particular task as being extremely important. Thus, the higher the score provided by a SME, the more critical the task was thought to be.

53. The tasks that were ranked as either a one or two are not particularly consequential. Dr. Davis explained that a value of greater then three meant that the officers thought that the particular task was critical. In Davis' validation study, jogging and running had values of 3.5; based upon the Delphi session, Dr. Davis' opinion was that these tasks were the most critical tasks.

54. The SMEs achieved the criticality rankings by consensus vote. Each SME would vote on a particular ranking and then the results would be entered into a laptop computer that would derive a mean value for the group. A group discussion among the SMEs would then occur as to why the SMEs voted as they did until a consensus was achieved. Normally, it would take two votes to achieve a consensus.

55. After computing the criticality rankings, Dr. Davis developed a scale regarding the frequency of performance of the tasks. A task which was performed daily was scored as a one; the performance of tasks that occurred weekly was a two; tasks done monthly were scored as a three; yearly tasks were scored a four; and a score of five indicated that the task was rarely performed.

56. Based upon a review of the scales used, Dr. Davis testified that there was a value of greater than five on swimming because the group basically did not do that task. In contrast, jogging received a score of 1.7 which means the SMEs expected jogging to take place almost on a daily basis.

57. Based on a review of the frequency and criticality rankings, Dr. Davis concluded that SEPTA officers walk with high frequency because the officers are predominantly foot-based. Dr. Davis also correctly concluded that SEPTA officers run more frequently than other police departments; he also found that they sprinted more often. In addition, Dr. Davis found SEPTA officers used a baton with more frequency than

in other jurisdictions. Overall, Dr. Davis assessed that the SEPTA officers are a more mobile and dynamic law enforcement group than most other law enforcement agencies.

58. After ascertaining the criticality and frequency rankings for the tasks on the Taxonomy, Dr. Davis had the SMEs define the perceived physical exertion for each task.

59. To determine what the perceived levels of physical exertion were for particular tasks, Dr. Davis asked the SMEs to identify, on the "Borg Scale," the perceived level of exertion for particular tasks. The Borg Scale is regarded as being scientifically authoritative and allows a person to identify the heart rate level that a particular physical task requires.

60. To determine the accuracy of the Borg Scale, Dr. Davis had previously tested the scale on Marine riflemen at the "Warfare Training Center." With respect to certain physical tasks, Dr. Davis compared the readings provided by the rifleman with the readings of the heart rates gathered from actual tests of the Marines. Dr. Davis found that the Marines' perceived ratings were incredibly close to the actual heart rate readings. This test confirmed Dr. Davis' understanding that the Borg Scale is an accurate, reliable and simple measure of physical exertion.

61. Based upon the SMEs' reading of exertion in SEPTA's jurisdiction relative to other police jurisdictions that Dr. Davis worked with prior to 1991, Dr. Davis concluded that most of the tasks engaged in by SEPTA officers are rated in a heart rate zone that would make the tasks aerobic in nature. In other words, the heart rate is elevated and there is an increased demand for oxygen to be supplied.

62. Dr. Davis testified that typical law enforcement officers simply do not engage in the type of activities with the same frequency as a SEPTA officer. The Court credits this testimony as being accurate. Indeed, the evidence introduced at trial establishes that SEPTA transit officers engage in physical activity more frequently than other law enforcement agencies.

63. After determining the physical exertion ratings for the particular tasks, Dr. Davis created physical dimension estimates for the particular tasks. To establish physical dimension estimates, one must place weight per measure on a task. In

essence, a distance, time and weight must be assigned to a particular task. For example, one can create physical dimension estimates for jogging by creating a jog where individuals are required to run one mile (distance) in seven minutes (time) while carrying 20 pounds of gear (weight).

64. To assign physical dimension estimates to the tasks in the Taxonomy, the SMEs gave Dr. Davis an estimate as to what they thought might be an expectation of physical dimension estimates for the tasks concerned.

65. In Dr. Davis' validation report, he set forth a series of tasks, such as a barrier surmount, long jump, stair-climbing, arrests simulation and distance run, along with the criticality and perceived exertion ratings interactive effect on these tasks.

66. Certain tasks from the original lists were combined because those tasks appeared to be logically linked together. For example, Dr. Davis combined sprinting with stair-climbing because the tasks are performed in close concert on the job at SEPTA. Dr. Davis also combined arrest simulation with defending one's self and effectuating arrests.

67. Dr. Davis subsequently estimated the SMEs' expectations regarding the appropriate time, weight and measure that should be placed on each task.

68. The expectations established that the SMEs would be standing for up to eight hours and walking for up to 6.3 hours. The expectation would be that a person might have to move a 142 pound person by himself. The SMEs also stated that 20 yards was a reasonable distance that they might be expected to move a 175 pound dead weight.

69. It was estimated that an officer could expect to run five flights of stairs. A reasonable expectation as to the height of a barrier that needed to be surmounted was 5.8 feet. The distance that an officer would be expected to jump over or cross was judged at 5.9 feet.

70. The SMEs stated that it was reasonable to expect them to have to run one mile in full gear in 11.78 minutes. Dr. Davis, however, rejected this information when creating the 1.5 mile run as a component of SEPTA's physical fitness test because the pace that the SMEs established was too low in Dr.

Davis' opinion. Dr. Davis believed that this physical dimension estimate was low because if such a pace was established as a test, it would require an aerobic capacity that almost any person could meet. Thus, if you were to use this estimate as a component of a physical abilities test, this component of the test would have no utility because almost any person could satisfy this minimal requirement. Based on Dr. Davis' experience and professional medical literature, Dr. Davis rejected this estimate as wholly unrealistic; the Court agrees with this assessment.

71. As the next step in his validation study, Dr. Davis attempted to determine the energy costs of performance of the physical tasks listed in his study.

72. Dr. Davis testified that all tasks performed by people have some nominal energy costs associated with them. For example, the aerobic capacity required to perform such activities as walking or standing is often measured by "METs" — the term MET is used interchangeably with the measurement of milliliters of oxygen per kilogram of body weight per minute or VO2 max. A MET is a multiple of the resting level of oxygen uptake, and one MET is resting level — it equals an oxygen uptake of 3.5 mL/kg/min which reflects the minimum energy required to maintain vital functions in the waking state. The amount of oxygen required to perform physical activity is evaluated in multiples of the resting metabolic rate. For example, a VO2 max of 42 mL/lg/min is equivalent to 12 METs (12 x 3.5 = 42). Dr. Davis testified that two METs would be doing twice resting — walking very slowly would be similar to two times the resting metabolic level, which is also called the "basal state."

73. Dr. Davis testified that a person's weight, the distance travelled and the time it takes to move over a distance are important factors in making energy cost calculations.

74. For example, to determine the energy cost to a SEPTA officer in moving from one point in space to another point in space, Dr. Davis had to consider the weight of an average SEPTA police officer — which is about 170 lbs — and the weight of the officer's load-bearing equipment. At the time of the Davis study in 1991, the extra weight added by the

equipment was about 12 pounds; this weight is now closer to 26 pounds. Indisputably, this extra weight slows an officer down. Moreover, the less a person weighs, the more this person would bear a disproportionate share of the "slowing down" caused by the extra weight because lighter persons do not have the mass to carry this extra weight. For example, 26 pounds of equipment will affect a 110 pound person more than a 170 pound person.

75. To calculate the energy cost for a person jogging at a pace of 200 meters per minute, the calculation is 2/10 of a milliliter of oxygen for each kilogram of the officer's body weight, plus the resting value, which equals 43.5 milliliters of oxygen to perform the task. This is the energy cost of running at a rate of 200 meters per minute.

76. Dr. Davis testified that although 43.5 mL/kg/min would be sufficient to move an officer at a rate of 200 meters per minute, this aerobic capacity would be insufficient for a SEPTA officer to perform his "job" under certain circumstances. In order to demonstrate under what circumstances 43.5 mL/kg/min would be insufficient for on officer to perform his duties, Dr. Davis described a scenario that SEPTA officers engage in frequently, that is, responding to officer assist or officer backup calls.

77. Dr. Davis explained that if an officer assist or officer backup call was reported and no train was available to take a SEPTA officer, who has an aerobic capacity of 43.5 mL/kg/min, to the next station where the officer requesting assistance was located, the "responding" SEPTA officer would have to make a decision as to whether she can get to the officer requesting assistance faster by foot or by waiting for the next train.

78. If the responding SEPTA officer chose to run to the next station, and ran at an 8-minute mile pace, this officer would arrive at the next station with some degree of energy reserve. Although the energy requirement to get the responding officer to the next station would be 43.5 mL/kg/min, this aerobic energy requirement is not the same as a person's maximal oxygen uptake. Instead, the person's maximal oxygen uptake is what is required to actually do the job, that is, to

engage in whatever activity is required of an officer when he arrives at the scene of the call. Dr. Davis testified that a SEPTA officer, under this scenario, would need a peak value of greater then 43.5 mL/kg/min in order to perform successfully.

79. Dr. Davis also testified that SEPTA officers have to surmount station steps once they get to the next station, which is going to require a higher energy reserve from the officer. The energy costs of stair-climbing is extremely high. Dr. Davis calculates that the aerobic capacity/energy consumption involved in ascending the typical flights of SEPTA stairs is 54 mL/kg/min. Thus, Dr. Davis testified that 54 mL/kg/min would be the expected aerobic capacity of a SEPTA officer who had to run up the typical flights of SEPTA stairs after running five to eight blocks.

80. In explaining how a person can undertake an activity by relying on their aerobic energy system, Dr. Davis used the term "pay as you go." This term means that as soon as one begins to exercise after being in a basal state, the demand for energy starts; and if the energy demand is below the person's maximal oxygen uptake, then the person can continue to exercise for a fairly substantial period of time.

81. Dr. Davis explained that a person with a high aerobic capacity will be able to take in oxygen, "turning it around and flushing out the lactate, which is the breakdown of the metabolites, and resynthesizing that lactate," enabling that individual to engage in physical activity for some period of time. For example, a marathon runner is able to run five-minute miles over a 26-mile distance due to his well-developed aerobic capacity system. Simply put, the person is able to supply himself with the energy needed to engage in a particular physical task by relying solely on his aerobic energy system.

82. Once the physical demands of an activity cannot be sustained by energy from the cardiovascular/aerobic system, a person is said to be hypoxic. When this occurs, physical activity comes to a halt because the person cannot supply the energy required to continue the physical activity. Thus, if a

person does not have a high aerobic capacity, this person will not be able to perform those physical activities that require a high aerobic capacity.

83. Dr. Davis does not agree with plaintiffs' expert, Dr. William McArdle, who indicated that SEPTA transit police officers were conducting and performing activities that were predominantly anaerobic.

84. As discussed above, aerobic capacity describes a person's maximum capacity for consuming oxygen during sustained physical activities. The body's ability to utilize oxygen for energy metabolism becomes important in sustaining physical activities like running, swimming, cycling and cross-country skiing.

85. Aerobic (*i.e.*, sustained, oxygen-utilizing) metabolic processes are to be distinguished from all-out exercise for short durations which are powered mainly by anaerobic chemical reactions which do not require molecular oxygen. Rapid anaerobic energy production maintains a high standard of performance in activities requiring all-out bursts of exercise such as sprinting in track and swimming, or repeated stop-and-go activities like soccer, basketball, volleyball, ice and field hockey, tennis and football. This quick, short-term energy production system — the anaerobic system — is used in "fight or flight" situations. The aerobic and anaerobic energy systems are separate energy systems, which use different types of muscle fibers and energy pathways. Anaerobic energy is generated largely by fast-twitch muscle fibers, which are fast-contracting and activated during intense change-of-pace and stop-and-go activities, as well as during all- out exercise that requires rapid, powerful movements. Aerobic energy, on the other hand, is generated largely by slow-twitch muscle fibers with a relatively slow speed of contraction compared to their fast-twitch counterparts. The primary role of the slow-twitch fibers is to sustain continuous endurance-type activities that require a steady rate of aerobic energy transfer, such as endurance running.

86. Because aerobic and anaerobic energy systems are distinct, a person with high aerobic capacity may not necessarily be able to perform well on tasks requiring short-term, pow-

erful bursts of energy. In sports competition, for example, elite endurance athletes with high levels of aerobic fitness generally do not excel in physical competition requiring short bursts of anaerobic energy and vice versa for elite sprint and high-power athletes.

87. Dr. McArdle testified that performance on SEPTA's 1.5 mile running test is almost totally influenced by one's aerobic capacity. In contrast, all-out runs for 3 minutes would be significantly affected by a person's short-term sprint or anaerobic capacity. Sprinting activities of less than 3 minutes are even more reliant upon anaerobic metabolic processes. According to Dr. McArdle, responding quickly to emergency situations like sprinting after a suspect, sprinting up stairs or running several blocks to assist a fellow officer requires short-term "bursts of energy" and call upon the body's anaerobic energy system. The same is true with respect to struggling with suspects and effectuating arrests.

88. Based on his belief that SEPTA officers typically run, at most, two to three blocks during the course of their duties rather than long distances, Dr. McArdle opined that SEPTA's 1.5 mile run does not measure the correct physical aspect of candidates, that is, a person's anaerobic capacity.

89. Although Dr. McArdle agrees that a 1.5 mile run is a useful field test for measuring aerobic capacity, he believes that the 1.5 mile run of SEPTA's application process tests for a physical ability — aerobic capacity — that is not required of SEPTA transit officers.

90. The Court, however, finds that Dr. McArdle's opinion is not supported by the record.

91. As an initial matter, the Court notes upon which Dr. McArdle never requested depositions to allow him to estimate the number of blocks that SEPTA transit officers run or the time it takes the officers to effectuate officer backups or officer assists on foot. Further, counsel for plaintiffs never supplied Dr. McArdle with depositions to allow him to estimate the number of blocks run or the time it takes SEPTA officers to effectuate officer backups or officer assists on foot.

92. In contrast to the dearth of information that Dr. McArdle based his opinion, the vast majority of evidence

introduced at trial indicates that SEPTA transit officers engage in runs or jogs on a daily basis that range anywhere from three to eight blocks and from one-quarter to one-fourth of a mile for periods ranging from three to ten minutes. This type of physical activity clearly entails a significant aerobic contribution.

93. Dr. McArdle actually acknowledged that the aerobic contribution would be significant in a number of actual experiences of SEPTA officers. For example, Dr. McArdle recognized that Experience 4 documented in Dr. Davis' validation report for SEPTA describes a five block run to arrest a perpetrator that would require a 50% to 60% aerobic contribution. Dr. McArdle admitted that the run in Experience 4 would require more than 50% to 60% aerobic contribution if the run was paced. Dr. McArdle further admitted that Experience 9 documented in Dr. Davis' validation report — a five-minute run to arrest a perpetrator — would require a major aerobic contribution of approximately 70%. Dr. McArdle admitted that the aerobic contribution would be even larger if Experience 9 was performed at a submaximal pace.

94. Further, Dr. McArdle acknowledged that the deposition testimony of many SEPTA officers describes running scenarios which would require significant contributions from the aerobic energy system. Under some of the running scenarios described, Dr. McArdle acknowledged that the aerobic contribution would be as high as 90%.

95. In addition, SEPTA officers do not engage in exercise on a maximal level, that is, they do not run as far as they can and as fast they can until they are totally incapable of doing any physical activity at the end of that exercise. SEPTA officers do run fast, but the best evidence suggests that SEPTA officers pace themselves when responding to officer assist or backup calls so that they will have some energy reserve when they arrive at the location to which they are responding.

96. Dr. McArdle defines paced-running efforts, *i.e.*, any effort that is not "all out", as submaximal. Dr. McArdle conceded that submaximal efforts increase the aerobic contribution and decrease the anaerobic contribution. Thus, Dr. McArdle implicitly concedes that the aerobic contribution to

a SEPTA officer's physical activity will rise in inverse proportion to the anaerobic contribution whenever that officer engages in submaximal activity.

97. Dr. McArdle has also written a book, titled *Exercise Physiology*, that contains a graphic depiction of aerobic and anaerobic contributions to maximal activity over time. Dr. McArdle's own graph demonstrates that at only ten seconds of maximal effort, the contribution is 10% aerobic, 90% anaerobic. Dr. McArdle's graph demonstrates that at only two minutes of maximal effort, the contribution is 50% aerobic and 50% anaerobic. Dr. McArdle's graph further establishes that at only four minutes of maximal effort, the contribution is 35% anaerobic, 65% aerobic.

98. Dr. McArdle also admitted that the more recent aerobic/anaerobic contribution charts of Jon Medbo and Izumi Tobatha published in the *Journal of Applied Physiology*, a respected journal in the field of physiology, demonstrate that the aerobic contribution for maximal running efforts occurs in greater percentages from as early as the first minute of running than that reported in Dr. McArdle's text book chart.

99. Based on the foregoing findings, the Court conservatively concludes that the relative contribution of aerobic capacity and anaerobic capacity at the two minute level is that 50% of the energy is supplied by the anaerobic energy system and 50% is supplied by the aerobic energy system.

100. The Court thus finds that it cannot credit Dr. McArdle's testimony that SEPTA is testing for the incorrect physical ability – aerobic capacity. Based on the evidence introduced at trial, especially Dr. McArdle's own published work and testimony, the Court finds that aerobic capacity is patently needed to be able to effectively perform many of the duties of a SEPTA transit officer; thus, SEPTA should be entitled to test for this physical ability.

101. The Court also credits Dr. Davis' testimony that pacing by SEPTA officers on officer assist calls will affect the relative contribution of aerobic metabolism versus anaerobic metabolism in that the set point at which an individual is going to run is going to be a function of their personal level of fit-

ness. The higher the fitness level of the individual, the greater the velocity that this person is going to be able to run from station to station.

102. Indisputably, the rate at which an officer performs an activity will be a function of the personal fitness level of that officer and that officer's work pace. The officer who has a high aerobic fitness level will have a greater energy reserve once she arrives at the location of an officer assist or backup call and is going to be able to do something more proficiently vis-a-vis the other officer with a low aerobic capacity who was trying to maintain a pace for which he cannot supply oxygen on an ongoing basis.

103. Moreover, the evidence establishes that the more aerobically fit an individual, the more quickly an individual will replenish their energy system whether the task performed requires aerobic or anaerobic energy. The "aerobic recovery pathway" is the process through which a person's energy system is replenished, enabling a person to continue to engage in physical activity; this replenishment process occurs in the mitochondria which are organelles inside a cell. Regardless of whether a task is aerobic or anaerobic, the payback mechanism is going to occur the same way. Because the energy required to continue a physical activity is going to be returned to the cells through the aerobic recovery pathway, regardless of whether the activity requires anaerobic or aerobic energy, it is undisputed that more aerobically fit individuals can replenish their energy system faster than less aerobically fit individuals.

104. The officer with greater aerobic capacity will be able to run faster and will be able to run for longer periods of time at a lower lactate level. Because more aerobically fit individuals run at a lower lactate level, the waste products of metabolism will be lower in this person's blood stream, rendering them more capable to perform the next series of events such as assisting another officer in a physical confrontation.

105. In addition, a high aerobic fitness level will clearly buffer the "follow-on demands," such as engaging in a difficult altercation or subduing a resisting arrestee, because this person is in much better condition physically than an individual

who has a lower aerobic capacity. Even if the follow-on demands are anaerobic, it is Dr. Davis' opinion that a higher aerobic capacity supports an individual's ability to carry out the anaerobic demands.

106. Dr. Davis concluded, based upon the validation study he did for SEPTA, that a SEPTA transit officer needs an aerobic capacity of 50 mL/kg/min to successfully perform a number of tasks.

107. Dr. Davis explained his findings to Ms. Pierce as to what he thought the job would require *i.e.*, that the job would require an aerobic capacity of 50 mL/kg/min. However, he also explained that such a standard would clearly have a draconian effect on the possibilities of women being hired as SEPTA officers.

108. Ms. Pierce specifically told Dr. Davis that she did not want the SEPTA police department to become the "boneyard" of the Philadelphia Police Department. Mr. Davis understood that Ms. Pierce was not concerned with having a standard that might be perceived as difficult for women to achieve; the job relatedness of the mission came first. In essence, SEPTA wanted to hire individuals who could perform the physical tasks required of a SEPTA officer regardless of whether this person was a man or woman; the Court finds that there certainly is nothing invidious about this goal.

109. Nevertheless, it was Dr. Davis' opinion that setting an aerobic capacity requirement in a range of 48 to 50 mL/kg/min would have an adverse effect on women because normative data demonstrates that there is a fairly substantial difference in terms of oxygen uptake and metabolism capabilities on the part of women as compared to men. Based on the normative data, Dr. Davis believed that a standard of 48 to 50 mL/kg/min would present a fairly substantial obstacle for women to seek employment with SEPTA.

110. Consequently, because Dr. Davis believed that the goals of SEPTA could be satisfied by using a 42.5 mL/kg/min standard for aerobic capacity, and because this standard would substantially reduce the adverse impact of a 50 mL/kg/min standard, Dr. Davis recommended to SEPTA that it set its aerobic capacity requirement at 42.5 mL/kg/min.

111. Dr. Davis felt that women could attain a standard of 42.5 mL/kg/min. Dr. Davis based this opinion on a project his company did for St. Paul, Minnesota, in which applicants for the fire department had to successfully run one and one-half miles in 11 minutes and 40 seconds. The aerobic capacity required to complete this run is 45 mL/kg/min. The outcome of the run was that out of the 705 individuals who applied for employment, 585 males and 120 females, 80% of the men passed and 76% of the women passed.

112. In addition to relying on the St. Paul data to support the aerobic capacity score of 42.5 mL/kg/min for SEPTA, Dr. Davis relied on previous test results of public service personnel that he tabulated during the 1980s. Dr. Davis administered these aerobic capacity tests to as many as 10,000 men and women between 1980 and 1991. Based upon this testing, Dr. Davis determined that 42.5 mL/kg/min is obtainable by women in sufficiently significant numbers to meet SEPTA's standard.

113. Prior to creating SEPTA's test, Dr. Davis also performed a study in Anne Arundel County, Maryland which supported an aerobic capacity requirement of 42.5 mL/kg/min for SEPTA. Dr. Davis tested 100 of Anne Arundel's incumbent officers on a battery of fitness tests and also on a job simulation test. There was a gym-based profile for all of the officers as well as a field test which included shooting a weapon, surmounting a barrier, a foot pursuit, a jump, a victim drag and an exit shooting scenario.

114. The data from this Anne Arundel County study showed that there were links between the physical fitness tests and the measures of fitness that were being tested, and the test also established a link between the gym-based fitness tests and the criterion tests. The measures of fitness that were tested were aerobic capacity and muscular strength measures such as grip strength, torso strength and muscular endurance.

115. After these tests were performed in Anne Arundel County, Dr. Davis examined the statistical relationship to establish whether the particular test would predict performance on the job tasks. The statistical procedure is called a "canonical correlation mobile regression analysis." Dr. Davis

testified that he observed statistical relationships between the tests and job tasks; in essence, passing performances on the test predict successful performance on the job.

116. Dr. Davis incorporates the Anne Arundel study into the SEPTA study where he discusses, at page 20, the "chi-squared analysis of time vs. aerobic capacity."

117. Dr. Davis, however, also conducted another job validation study for a sheriff's department in Florida in which he could not establish that passing performance on a 1.5 mile run correlated with good performance. Nevertheless, there was no evidence introduced at trial that the tasks performed by these sheriffs were similar to the tasks performed by officers at SEPTA. Thus, the Florida study has no impact on whether the 42.5 mL/kg/min requirement for SEPTA officers is valid.

118. Dr. Davis started the project at SEPTA with the objective of designing a criterion task test that could be administrated to SEPTA officers. These criterion tasks are essential functions that a police officer might be expected to do on the job. The criterion tasks that Dr. Davis was going to create for SEPTA is known as a work sample; these tasks included a barrier surmount, long jump, stair climbing, arrest simulation and a distance run.

119. The distance run component of the test would be used to determine whether the applicant had the required minimal level of aerobic capacity that Dr. Davis had previously identified as being necessary to perform the job of SEPTA transit officer.

120. Because Dr. Davis wanted to test for an aerobic capacity of 42.5 mL/kg/min, Dr. Davis suggested that SEPTA implement a distance running test whereby applicants would be required to run 1.5 miles in 12 minutes or less. Dr. Davis suggested this distance and time because if an applicant could complete the run in 12 minutes or less, it could be concluded that the successful applicant had an aerobic capacity of at least 42.5 mL/kg/min.

121. Although there was some evidence introduced at trial that a one and one-half mile run equates to an aerobic capacity of 43.5 mL/kg/min, the validity of Dr. Davis' suggested aerobic capacity requirement of 42.5 mL/kg/min is not

affected because the standard error for determining whether a 1.5 mile run in 12 minutes measures an aerobic capacity of 42.5 or 43.5 mL/kg/min is about two or three mL/kg/min. Thus, Dr. Davis' suggested aerobic capacity of 42.5 mL/kg/min is an appropriate measure in light of the standard error.

122. Because Dr. Davis believes that a 1.5 mile run in 12 minutes identifies persons who possess a reasonable level of stamina to perform the essential elements of a job, he suggested to SEPTA that it use the distance run as a "front end screen," *i.e.*, SEPTA should require the applicants to run the 1.5 mile course as the first step in a physical fitness test.

123. Dr. Davis adopted the 1.5 mile run versus laboratory testing because laboratory testing would cost hundreds of thousands of dollars in equipment; whereas, SEPTA could administer the run outside for considerably less money. Second, Dr. Davis believes that almost all of the applicants would be familiar with a run conducted outside.

124. Dr. Davis did not use a shorter distance than 1.5 miles for several reasons. First, in order to appropriately test for this construct or dimension of fitness (aerobic capacity), the test has to be approximately twelve minutes. Second, if the length is changed to a shorter distance, the dynamics of the metabolic system change, thus distorting the ability to estimate aerobic capacity. Third, shorter runs, which rely more on anaerobic capacity, dramatically inflate the differences between men and women and would thus further disadvantage women for consideration for employment.

125. Dr. Davis testified that it was not his understanding that SEPTA transit officers are running 1.5 miles in the course of their duties. Nevertheless, he still suggested that SEPTA should use the 1.5 mile run as part of its physical fitness test because the run was not being used to simulate an actual job event, rather it was being used to test the construct of aerobic capacity.[4]

---

4. Dr. Davis is familiar with the concept of construct validation. Indeed, Dr. Davis testified that he has read the *Uniform Guidelines* and the *Principles for the Validation and Use of Personnel Selection Procedures* (1987) ("SIOP Principles"). The SIOP Principles set forth the profession's standards for the choice, development, evaluation and use of personnel selection

126. In Dr. Davis' profession of exercise physiology, a construct can be interpreted as a dimension of fitness. Fitness is defined into five components: cardiovascular fitness or stamina; muscular strength; muscular power; explosive strength; and muscular endurance. In addition to these five dimensions, there exist a number of motor skills that can be tested.

127. The construct that the 1.5 mile run was designed to measure is stamina, the ability to take up and utilize oxygen, *i.e.*, aerobic capacity.

128. In the course of creating the physical abilities test for SEPTA, Dr. Davis was able to link aerobic capacity to the specific critical tasks that he observed SEPTA officers doing on the job. Dr. Davis testified that the link is common sensical in that every job task analysis that has ever been done for any reasonably proactive law enforcement organization finds that running is a critical and essential task. Also, statistical manipulations have been established showing that there exists a correlation between police officer performance and a 1.5 mile run. Dr. Davis testified that these statistical findings have repeatedly proven that which he believes is obvious, that is, "if you have a good cardiovascular system you can do the job, if you have a big cardiovascular system you can do more of the job."

129. In sum, the Court finds that Dr. Davis demonstrated that an aerobic capacity of 42.5 mL/kg/min is necessary to successfully perform the functions of a SEPTA transit officer.

130. With respect to the remaining elements of Dr. Davis' proposed criterion task test, Davis discovered that SEPTA did not have the space to set up of a permanent criterion-task test course. In addition, he was informed by SEPTA that SEPTA had a contract with the Benjamin Franklin Clinic, which is associated with the Pennsylvania Hospital in Philadelphia, under which its officers were already undergoing a physical fitness program. Thus, to maximize its money, SEPTA asked

---

NOTES *(Continued)*

procedures. The *Uniform Guidelines* establish standards for assessing the job-relatedness or validity of employee selection devices.

Dr. Davis to create a "surrogate" for the criterion-task test that they could administer at the Benjamin Franklin Clinic.

131. Dr. Davis recommended certain muscular fitness measures which would test for muscular strength and endurance — physical abilities that are needed to successfully perform as a SEPTA transit officer. In this regard, Dr. Davis recommended a grip strength test, bench press, pull-ups, push-ups and sit-ups. In the Anne Arundel study, all of these tests were found to be predictive of successful performance on police work.

132. Because flexibility does not predict job performance, Dr. Davis did not create a test that would test for this fitness measure.

133. Although Dr. Davis did not set the levels required to be met for each component of the physical fitness test to include men, Dr. Davis did set the levels at a point where he felt that the goals of SEPTA could be achieved and that women would not be unreasonably excluded. Based on his vast experience in creating physical fitness tests, Dr. Davis concluded that each fitness test was achievable by women.

134. The Court, contrary to plaintiffs' contention, finds that Dr. Davis was a credible witness and was not biased against women. His testimony, especially regarding his validation study, was both credible and objective.

135. Finally, the Court notes that Dr. Davis testified that SEPTA could have engaged in "rank-ordering" hiring — hiring from a list a successful test-takers from top to bottom until all positions are filled.

3. The Physical Fitness Test

136. Based on Dr. Davis' recommendation, SEPTA adopted a test which consisted of the following components with the following scores for passing: (1) a 1.5 mile run that must be completed in 12 minutes or less; (2) bench press — 5 repetitions of 115 pounds; (3) grip strength — 100 pounds as measured on a dynamometer with the dominant hand; (4) pull-up — 1 pull-up, palms away, from a dead hang, elbows

flexed to allow the chin to clear the bar; (5) push-ups – 30 repetitions of "military style" push-ups; and (6) sit-ups – 45 repetitions in two minutes.

137. The components of the physical fitness test other than the 1.5 mile run are commonly referred to as the "gym-based components."

138. SEPTA's physical fitness test also originally included a body fat measurement. Men were required to have less than or equal to 29% body fat; women were required to have less than or equal to 22% body fat.

139. Starting in 1991, applicants participating in SEPTA's physical fitness test were first administered the 1.5 mile run component. If an applicant failed to complete the 1.5 mile run in the required time, he or she was disqualified from the selection process and not permitted to participate in the gym-based components of the physical fitness test.

140. Applicants who passed the 1.5 mile run were invited to participate in the gym-based components. These components were administered by the Benjamin Franklin Clinic.

141. SEPTA administered the above physical fitness test, including the 1.5 mile run, to transit police officer applicants in July 1991 and October 1993. In 1991, the run was administered on the exercise field at Temple University. In 1993, the run was administered in Fairmount Park. There was also evidence admitted at trial that indicates that the physical fitness test may have been administered to some women in 1992.

142. The 1992 female applicants were provided with only a few days notice of the 1.5 mile run requirement between the time they took the written test and the time they were administered the running test. Each of the approximately five to six female applicants who took the 1992 running test failed. One of these applicants was told that she could return for a retest the following day.

143. SEPTA changed its physical fitness test for transit police officer applicants in or around late 1995 or early 1996. Although it retained the 1.5 mile run component, SEPTA abandoned the gym-based components and replaced them with the following four components: (1) turnstile jump – vault over a SEPTA turnstile; (2) barrier surmount – scale a

six foot chain link fence; (3) dummy drag — drag a 170-pound dummy 100 feet; and (4) weapon test-fire — squeeze trigger of 9 millimeter weapon twenty times with each hand. These components are commonly referred to as the "criterion tests." According to SEPTA, the criterion tests measure the same fitness constructs as the gym-based components.

144. At the time SEPTA abandoned the gym-based components and replaced them with the criterion tests, SEPTA was aware that administrative complaints had been filed by the *Lanning* plaintiffs with the Pennsylvania Human Rights Commission in April 1994.

145. SEPTA administered this new physical fitness test — including both the criterion tests and the 1.5 mile run — to transit police officer applicants on March 16, 1996. The 1.5 mile run was again administered at Fairmount Park. The criterion task tests were administered at the subway station at Broad and Pattison Streets in South Philadelphia. All applicants who took the criterion task tests in 1996 passed them.

146. Since 1991, SEPTA's selection procedure for hiring transit police officers has consisted of a written examination (graded pass/fail); the physical fitness test described above, involving the 12 minute, 1.5 mile run and either the gym-based tests or the criterion tests (graded pass/fail); an interview conducted by a panel of officials; an interview with the Chief of Police; a background investigation including a polygraph; and a medical examination including a drug screen.

147. Applicants passing the written examination are invited to participate in the physical fitness test as described above. Applicants who pass the physical fitness test are invited to participate in an oral interview conducted by a panel of three SEPTA officials who ask a standard set of six questions to all applicants. Each interviewer ranks the applicant's answer to each of the six questions. The applicant is then assigned an overall numerical score consisting of the total of the scores given by each of the three interviewers to the applicant's answer to each of the six standard questions.

148. Applicants are placed on an eligibility list in rank order based solely on their overall score on the panel interview. Scores on the written and physical fitness tests have no

bearing on an applicant's overall score or ranking on the eligibility list. Thus, the candidate with the highest score on the physical fitness test could be ranked last on the eligibility list; conversely, the candidate with the lowest passing score on the physical fitness test could be ranked first on the eligibility list. However, this result is irrelevant for the purposes of this case because any person who has passed the physical fitness test has demonstrated the ability to perform successfully on those tasks required of SEPTA officers. The applicant's overall score is not included on the eligibility list. Rather, the names are listed in rank order starting with the name of the top scorer.

149. When SEPTA is ready to fill a transit police officer vacancy, the Chief of Police conducts a personal interview with a number of applicants from the current eligibility list. The number of applicants interviewed depends on the number of vacancies to be filled. Applicants are called for interviews with the Chief of Police in rank order from the eligibility list. Absent some concern raised during the personal interview, the Chief of Police makes offers of employment to applicants in rank order from the eligibility list based on the number of vacancies that are needed to be filled. The medical examination and background investigation are completed before an offer of employment is made.

150. Eligibility lists remain effective until they expire or are exhausted, but in no case does the eligibility list remain in effect longer than three years. In some instances, offers of employment are not made until many months — and in some cases, two and a half years — after the physical fitness test is administered.

151. Applicants are not instructed or required to maintain their physical fitness or to engage in any type of exercise regimen while they are on the eligibility list. No retesting is done to determine whether those on the eligibility list can still meet the physical fitness test at the time an offer of employment is made. It is after an offer of employment is made that the applicant enters the Philadelphia Police Academy ("Academy"), a fully accredited state police training academy.

D. The Named Plaintiffs, Class Members, and Test Passers

152. At the time of their applications in 1993, each of the individual named plaintiffs had the minimum qualifications for the position of transit police officer.

1. Plaintiff Catherine Natsu Lanning

153. Plaintiff Catherine Natsu Lanning is a female resident of Fairless Hills in Pennsylvania. She was 26 years old in 1993 when she took the 1.5 mile run component of the physical fitness test for SEPTA transit officer applicants. She was also a United States citizen and held a valid driver's license.

154. Ms. Lanning graduated from Montgomery County Community College Police Training Academy in June 1993. This academy is a fully accredited state police academy. She was valedictorian of her class, having a grade point average of 98% and having scored an overall 95% on the physical fitness tests.

155. Ms. Lanning has been employed as a police officer at the University of Pennsylvania since May 1994. Since October 1995, she has been assigned to the bike unit, where she patrols West Philadelphia by bicycle for approximately three (3) out of five (5) shifts per week. Ms. Lanning was recently selected to serve as one of the first officers on the University of Pennsylvania's elite tactical bike patrol unit to focus on reducing and deterring serious crime on the University of Pennsylvania's campus. Ms. Lanning's duties have included patrolling SEPTA stops in her jurisdiction and providing backup assistance to SEPTA officers. Ms. Lanning has received commendations for her heroic performance at the University of Pennsylvania.

156. Ms. Lanning is certified by the Commonwealth of Pennsylvania under Act 120 ("Act 120") as having met state mandated physical fitness, psychological, academic and training requirements for police officers. Act 120 certification is required of all municipal police officers in Pennsylvania.

157. Ms. Lanning admits to receiving a SEPTA "pamphlet" during the application period that provided guidance on how to train to meet the 1.5 mile running test. Nevertheless, Lanning made no effort to follow the instructions on how

to improve her running time that were contained in the pamphlet that SEPTA sent her. When she actually participated in SEPTA's 1993 applicant running test, Lanning ran some portion of the course with her hands in her pockets.

2. Plaintiff Altovise Love

158. Plaintiff Love is a female resident of Pennsylvania. She was 23 years old in 1993 when she took the 1.5 mile run component of the physical fitness test. She was also a United States citizen and held a valid driver's license.

159. Ms. Love graduated from Northeast High School and attended classes at Community College of Philadelphia. Ms. Love also graduated from the Academy.

160. Ms. Love has been employed as a police officer for the Philadelphia Police Department since October 1994. Ms. Love has worked on foot and car patrol for the Philadelphia Police Department and is currently assigned to the 15th Police District, a high crime area. Ms. Love's duties regularly include responding to calls for crime occurring on SEPTA property or providing backup assistance to SEPTA officers. Ms. Love is currently on the list for promotion to a detective position with the Philadelphia Police Department.

161. Ms. Love is certified by the Commonwealth of Pennsylvania under Act 120 as having met state mandated physical fitness, psychological, academic and training requirements for police officers.

162. Ms. Love admits that she did not prepare for SEPTA's 1.5 mile applicant run.

3. Plaintiff Belinda Kelly Dodson

163. Plaintiff Dodson was a female resident of Pennsylvania and was 30 years old in 1993 when she took the 1.5 mile run component of the physical fitness test for SEPTA police officer applicants. She was also a United States citizen and held a valid driver's license.

164. Ms. Dodson has an associate's degree in forensic science and police science from New River Community College in Dublin, Virginia. She is currently pursing her bachelor's degree in law enforcement at George Mason University.

165. Ms. Dodson has successfully passed the Virginia Commonwealth physical fitness and academic requirements necessary for Virginia Commonwealth certification as a police officer.

166. Ms. Dodson has over ten years of law enforcement and other related experience. This includes working as a sheriff for the Fairfax County Sheriff's Department in Virginia from 1986 to 1989, as a police officer at George Mason University from 1989 to 1992, and as a police officer for Swarthmore College Public Safety Department. Ms. Dodson received a commendation for her heroic work as a police officer at George Mason University. Her work as a police officer has included both foot and car patrol duties.

167. Ms. Dodson was appointed by the Commonwealth of Pennsylvania as a private police officer.

168. Although Ms. Dodson was running between two and three times a week during the period of her application to SEPTA, she admits that she never changed her exercise routine upon learning of SEPTA's running test. She also never timed herself on practice runs prior to taking SEPTA's 1993 applicant running test. Ms. Dodson further stated that her run in SEPTA's 1993 1.5 mile applicant running test was appropriately characterized as a "slow jog." After failing SEPTA's 1.5 mile applicant running test in 1993, Ms. Dodson never reapplied to SEPTA.

4. Plaintiff Denise Dougherty

169. Plaintiff Dougherty is a female resident of Philadelphia, Pennsylvania. She was 22 years old in 1993 when she took the 1.5 mile run component of the physical fitness test for SEPTA. She was also a United States citizen and held a valid driver's license.

170. Ms. Dougherty has completed three years of course work in Criminal Justice at Temple University. Since April 1996, Ms. Dougherty has been employed as an administrator at Hear Now in Philadelphia.

171. Prior to the running test, Ms. Dougherty believed SEPTA's running test was a reasonable test and that she could successfully complete the test without training. Ms. Dougherty

did nothing to prepare for the running test. Further, Ms. Dougherty admits to walking during portions of SEPTA's 1.5 mile running test.

5. Plaintiff Lynne Zirilli

172. Plaintiff Zirilli (formerly Lynne Carapucci) is a female resident of Philadelphia, Pennsylvania. She was 24 years old in 1993 when she took the 1.5 mile run component of the physical fitness test for SEPTA police officer applicants. She was also a United States citizen and held a valid driver's license.

173. Ms. Zirilli graduated from St. Maria Goretti High School in Philadelphia. Ms. Zirilli was hired as a police officer by the Philadelphia Police Department. In December 1997, she graduated from the Academy after successfully passing all physical and academic requirements. She is currently a police officer in the 3rd District, where her duties include routine checks of SEPTA property.

174. Ms. Zirilli is certified by the Commonwealth of Pennsylvania under Act 120 as having met state mandated physical fitness, psychological, academic and training requirements for police officers.

175. In 1993, Ms. Zirilli had never attempted to run on her own. When Zirilli learned of SEPTA's 1.5 mile running test, she did nothing to prepare for the run. Ms. Zirilli admits that she walked during various portions of SEPTA's 1993 1.5 mile applicant running test.

6. Class Member Kim French

176. Plaintiff Kim French is currently employed by the Philadelphia Police Department; Ms. French was hired in 1995.

177. Ms. French initially applied for a position with SEPTA in 1993. After passing the written test portion of SEPTA's application process, Ms. French ceased pursuing her application due to the fact that she was pregnant.

178. Ms. French, however, reapplied to SEPTA in 1996. After passing the written test portion of SEPTA's application process, Ms. French participated in SEPTA's 1996 1.5 mile applicant run, which she failed. Although Ms. French testified

that she rides a stationery bike and walks for exercise. Ms. French admits that she did nothing further to prepare for the 1.5 mile run. Ms. French concedes that she could train to run 1.5 miles in 12 minutes.

7. 1992 Test Taker - Dawn Kennedy

179. Dawn Kennedy is currently employed by the University of Pennsylvania Police Department and is assigned to the Bicycle Patrol Unit. In her current employment, Ms. French has dual jurisdiction over SEPTA property in some geographic areas (areas on the University of Pennsylvania campus); in effect, she patrols some of the same areas that SEPTA police officers patrol.

180. Although SEPTA contends it only accepted applicants for possible employment in 1991, 1993 and 1996, administering the 1.5 mile run during these application periods, Ms. Kennedy testified that she applied to SEPTA in 1992 for a position as a transit officer; the Court finds Ms. Kennedy's testimony to be credible.

181. After passing the written examination portion of SEPTA's application process, Ms. French was asked to participate in a 1.5 mile run which had to be completed in 12 minutes or less; SEPTA informed Ms. French that the run would be held in two days.

182. At the time Ms. French was informed by SEPTA that she had to participate in this run, Ms. French was training for an identical run in the Delaware County Municipal Police Academy. In order to train for the run, Ms. French, on the intervening day between the notice of the run and the actual run, went to a track and ran two miles in preparation for the SEPTA run.

183. On the day of the run, it was raining and cold. Ms. French ran with five or six other females. After the run, Ms. French was informed that she and the other runners failed the test.

184. On the same evening of the run, Ms. French was contacted by SEPTA and asked to participate in another run on the following day. Ms. French could not attend this run because of a prior appointment; she never reapplied to SEPTA.

8. SEPTA Officer Bernadette Rodier

185. Prior to hire as a SEPTA transit officer, Officer Bernadette Rodier spent one year working as a sales clerk for a uniform store and fifteen years as a waitress for a Denny's Restaurant.

186. In the summer or fall of 1996, Officer Rodier read an advertisement for the position of SEPTA transit officer that informed her that she would be required to run 1.5 miles in 12 minutes or less.

187. To prepare for SEPTA's running test, Officer Rodier ran outside and on a treadmill a few times a week, and she would time herself once a week to note her progress and to ensure herself that she would be able to run 1.5 miles in 12 minutes or less. She timed herself every Sunday from the time she started training until she took the test.

188. After passing the written portion of the SEPTA test, Officer Rodier received a pamphlet from SEPTA which contained suggestions on how to train for the physical fitness test; she received this pamphlet two to three months before the administration of the run. Officer Rodier followed these suggestions contained in the pamphlet.

189. After successfully passing the running portion of SEPTA's test, Officer Rodier took the remaining portion of the physical fitness test, which she passed.

190. Officer Rodier graduated from the Academy on July 15, 1997, and she spent an additional two weeks training. In the beginning of August 1997, Officer Rodier was placed on street patrol. Based on her experience as a SEPTA officer, Officer Rodier does not believe that the Academy provided her with sufficient training for her job as a SEPTA officer.

9. SEPTA Officer Margaret Gerlach

191. SEPTA Officer Margaret Gerlach graduated from high school in 1986. During and after high school, Officer Gerlach worked at Thrift Drug for two years. Officer Gerlach then started and ran her own cleaning service for two years. Subsequent to running a cleaning service for two years,

Officer Gerlach worked as a merchandiser for Nabisco. She then spent two years working for the University of Pennsylvania.

192. Officer Gerlach saw an advertisement in a newspaper for the position of SEPTA transit officer. From this advertisement, Officer Gerlach became aware that she was going to be required to run 1.5 miles in 12 minutes or less.

193. Prior to taking SEPTA's running test, Officer Gerlach received a pamphlet that instructed applicants as to how to prepare and train for the running test. Officer Gerlach followed a few of the suggestions contained in this pamphlet. To prepare for the running test, Officer Gerlach measured out one and one-half miles on a track and ran that course. Officer Gerlach ran 1.5 miles three times per week, always timing herself; she also went to the gym and used a stair climber, treadmill and bicycle to train for the run and weight trained to improve her upper body strength.

194. Officer Gerlach passed the running test and the subsequent physical fitness tests. She was eventually hired as a SEPTA transit officer.

10. SEPTA Officer Nicole Heppard

195. SEPTA Officer Nicole Heppard attended Pennsylvania State University and graduated with a criminal justice degree in 1992. She previously worked as a loss prevention detective for Strawbridge & Clothier and for the Sports Authority.

196. Officer Heppard saw SEPTA's advertisement for the position of SEPTA transit officer in the *Philadelphia Inquirer* in November 1996. The advertisement stated that applicants would have to run 1.5 miles in 12 minutes or less.

197. Officer Heppard took and passed the written test portion of SEPTA's application process; she was then invited to participate in the running portion. Prior to the run, Officer Heppard received a pamphlet from SEPTA, explaining how to train for the running test. At the time she received this information, Officer Heppard did not regularly exercise. Nevertheless, Officer Heppard began to prepare for the run approximately one month before the run in response to

receiving the training information. Officer Heppard began to run approximately two to four times per week and subsequently noticed that she was able to run farther each time she ran.

198. Officer Heppard passed the running portion of SEPTA's test and then passed the muscular strength and endurance portions of the test.

199. Officer Heppard currently works out to prepare for her incumbent physical fitness test. In this regard, she trains about two to three hours per week. Officer Heppard can currently pass the pull-ups, push-ups and sit-ups portions of the test. Officer Heppard has not been able to satisfy the bench press portion of the test.

11. Former SEPTA Officer Bridget McCarthy Poggi

200. Officer Bridget McCarthy Poggi is currently employed by the Springfield Police Department as a patrol officer. Officer Poggi was formerly employed by SEPTA as a transit officer prior to her employment with the Springfield Police Department.

201. Officer Poggi took SEPTA's 1.5 mile running test on the track at Temple University and successfully passed the test. She subsequently passed the muscular strength and endurance portions of the SEPTA application process.

202. To train for the run, Officer Poggi ran approximately five times per week and lifted weights approximately three to five times per week.

12. SEPTA Officer Tracy Thomas

203. SEPTA Officer Tracy Thomas was hired by SEPTA as a transit officer in 1991. Officer Thomas testified that she was hired in 1991 despite her failure on the 1.5 mile running test and the gym-based muscular strength and endurance test. Like the class members herein, Officer Thomas never timed herself on a 1.5 mile run prior to taking SEPTA's running test.

204. Like the class representatives, Officer Thomas admits that she was not regularly exercising at the time of the

run in 1991. Officer Thomas admits that she would have passed the running test if she was regularly running at the time of the test.

205. Officer Thomas has been able to train from an aerobic capacity of 33 mL/kg/min to 42 mL/kg/min. Officer Thomas credits this increase to her training and SEPTA's incumbent physical fitness testing program.

206. In sum, the female applicants who failed SEPTA's 1.5 mile running test in 1993 and 1996 all demonstrated a cavalier attitude toward the position by not preparing or training for the running test.

207. In contrast, the four female witnesses who passed the running test, regardless of their varying fitness backgrounds, all specifically prepared and trained for the running test to increase or ensure their chance for success.

E. Number of Women Among SEPTA's Ranks

208. SEPTA has an extremely low number of women among its sworn ranks. As of July 1997, SEPTA's sworn personnel consisted of a Chief, one Deputy Chief, 3 Captains, 11 Lieutenants, 28 Sergeants, and 190 patrol officers. Of these 234 sworn employees, there is only 1 female Lieutenant, 1 female Sergeant, and 14 female patrol officers.

F. Adverse Impact of the 1.5 Mile Running Test

209. SEPTA admits to the information contained in the following chart with respect to its administrations of the 12 minute, 1.5 mile running test to transit police officer applicants:

	1991	1993	1996	TOTAL
Number of Female Test-Takers	23	28	32	83
Number of Female Passers	6	1	3	10
Female Pass Rate	26.1%	3.6%	9.4%	12.1%
Number of Male Test-Takers	332	412	336	1080
Number of Male Passers	227	197	219	643
Male Pass Rate	68.4%	47.8%	65.2%	59.5%
Number of Standard Deviations	**2.42**	**3.38**	**3.88**	**5.56**
p-value	1/10,000	1/10,000	1/100,000	1/100,000

210. The row entitled "number of standard deviations" shows the disparity between the pass rate for male and female applicants as measured by the formula set forth in *Hazelwood School District v. United States*, 433 U.S. 299, 308, n.14, 97 S. Ct. 2736, 2742 n.14, 53 L. Ed. 2d 768 (1977) and *Castaneda v. Partida*, 430 U.S. 482, 496-97 & n.17, 97 S. Ct. 1272, 1281-82 & n.17, 51 L. Ed. 2d 498 (1977) (hereinafter "*Hazelwood* formula").

211. The disparities between the pass rates for male and female applicants as measured by the *Hazelwood* formula are all statistically significant at the .05 level, *i.e.*, the likelihood that the disparities can be accounted for by chance is less than 5 in 100.

212. The row entitled "p-value" is calculated using the Fisher exact 2-tail formula.

213. The results of the administrations of the 1.5 mile run in 1993 and 1996 (the period of time covered by the Lanning class) are set forth in the following chart:

	1993 and 1996
Number of Female Test-Takers	60
Number of Female Passers	4
Female Pass Rate	6.7%
Number of Male Test-Takers	748
Number of Male Passers	416
Male Pass Rate	55.6%
Number of Standard Deviations	**5.06**
p-value	1/100,000

214. In a June 24, 1996 memorandum, SEPTA's affirmative action officer, Judy Hirsch, stated that with respect to the 1996 physical fitness test for transit police officer applicants:

> A standard deviation analysis found the difference in the run pass rates between males and females to be grossly significant (5.9 standard deviations).

215. Although SEPTA has never admitted to or provided evidence about the number of women who took and failed the 12 minute, 1.5 mile running test in 1992, at least five female applicants took and failed the 1.5 mile running test during a test administration in 1992.

216. Thus, the disparate impact of SEPTA's 1.5 mile running test is slightly more pronounced than the above statistics reflect.

217. In addition to the empirical evidence in this case, research in the field of exercise physiology establishes that setting a cutoff score of 12 minutes on a 1.5 mile running test will have an adverse impact on women.

218. Scientific studies show that males score higher on tests of V02 max and endurance performance than their female counterparts due to physiological differences between men and women. This result is attributable to the well-documented sex differences in body composition and hemoglobin, the iron-containing compound in the blood responsible for oxygen transport because men have more muscle mass and less fat per unit of body weight than women. The most important factor determining one's capacity for oxygen consumption during exercise is the quantity of muscle mass a person possesses; this is because the site of aerobic metabolism occurs in the active muscles. It is partially because of this difference in the amount of potentially active muscle mass during exercise that men consistently score higher in VO2 max tests like the 1.5-mile run test administered by SEPTA.

219. Data from the Institute For Aerobics Research in Dallas, Texas (the "Cooper Institute") indicates that requiring men and women to run 1.5 miles in 12 minutes has an adverse effect on females. Based on studies of approximately 40,000 American men and women, the Cooper Institute has developed normative standards for determining the physical fitness of men and women of different ages on a range of fitness items. According to the Cooper Institute data, approximately 47% of men aged 20 to 29 years in the general population can achieve a 1.5 mile run time in 12 minutes. In contrast, only 12% of women in this age category can achieve this time. However, evidence introduced at trial indicates that the data produced by the Cooper Institute may not be reliable. Specifically, there was testimony at trial that these normative standards for determining the physical fitness of women may not be representative of all American women because the Cooper Institute used a sample of predominantly white women of

higher socioeconomic status who visit the Cooper Institute for specific medical reasons. Consequently, the Cooper Institute's normative standards for women may be only representative of a certain cross-section of American women and not representative of all American women. Indeed, Steven Blair, the current Director of the Cooper Institute, has recently suggested that the Cooper Institute will conduct a new national survey to measure aerobic capacity because of the limitations on the current normative standards published by the Cooper Institute.

220. At all times relevant to this litigation, SEPTA was aware of the disparate impact upon women caused by its 12 minute 1.5 mile running test. Nevertheless, SEPTA never undertook any study to determine whether alternative tests existed which would have less of an adverse impact on women.

221. Plaintiffs' expert, Sheldon Zedeck, Ph.D., who testified in this case as an expert in industrial and organizational psychology, test development and test validation, created a physical fitness test (an applicant physical fitness test for the San Francisco, California Fire Department) which has an adverse impact on women. Importantly, Dr. Zedeck testified that he has never returned to San Francisco to search for or create a new test that has less of an adverse impact on women. Dr. Zedeck also contends that he has not violated any professional standards by failing to search for alternative tests that may have less of an adverse impact on women.

G. Adverse Impact of Gym-Based Components

222. SEPTA claims that all four of the female applicants who have taken the gym-based components of the test (1 in 1993 and 3 in 1996) have passed these components and that therefore the United States cannot establish the adverse impact of the gym-based components.

223. However, the Court finds that 28 female applicants took and failed the gym-based components of SEPTA's physical fitness test in 1991, thereby refuting SEPTA's argument that every woman who has taken the gym-based components has passed them. Nevertheless, there is no data available to compare the pass rates of male and female applicants on the

gym-based components. Moreover, not one of the 28 female applicants passed the running portion of SEPTA's test; thus, they were ineligible to be hired as SEPTA officers.

224. The government's witness, Dr. McArdle, testified that there are studies in the field of exercise physiology showing that men, on average, score higher on tests of upper body strength. Due to physiological differences between men and women in the quantity of muscle mass and its distribution on the body, scientific research indicates that females typically have only about 50% of the upper body strength of male counterparts compared to 70% of the leg strength of males.

225. SEPTA's own test developer, Dr. Davis, acknowledged these differences and admitted that each of the muscular strength and endurance components of SEPTA's "gym-based" test (*i.e.*, bench press, push-up, sit-up, pull-up and grip strength) would have an adverse impact on women. However, as noted previously, Dr. Davis also acknowledged that women can train to meet and pass the physical fitness tests administered to SEPTA.

226. The sit-up, bench press and grip strength items are common to many physical fitness test batteries. When determining the "fitness" of men and women with these tests, some professionals in the field of exercise physiology recognize sex differences in physical performance capacity and evaluate test scores based on sex-specific standards, as is the case for the Cooper Institute normative standards.

227. Plaintiffs' expert exercise physiologist, Dr. McArdle, testified that (a) the requirement of SEPTA's physical fitness test that transit police officer applicants bench press 115 pounds for five repetitions has an adverse impact against females and (b) the requirement of SEPTA's physical fitness test that transit police officer applicants complete one pull-up from a dead hang also has an adverse impact against females.

228. Dr. Davis also acknowledged that women are at a distinct disadvantage with respect to performance on pull-up tests. He testified that the difference between men and women is dramatic: men outperform women by at least 500% and sometimes over 1000% on pull-up tests.

229. There was evidence admitted at trial that suggests that SEPTA's requirement that transit police officer applicants complete 30 military style push-ups would have an adverse impact against females. According to a database collected by the United States Army on the fitness of Army trainees, the strongest female Army recruits could perform a maximum of only 18 push-ups. However, there was no evidence introduced at trial that indicated whether these Army recruits trained before they took the test.

230. There was other evidence introduced at trial that the requirement of SEPTA's physical fitness test that transit police officer applicants complete 45 sit-ups in two minutes has an adverse impact against females, as does SEPTA's requirement that transit police officer applicants demonstrate 100 pounds of grip strength in the dominant hand as measured by a dynamometer.

H. Incumbent Officers

1. Physical Fitness Testing of Incumbent Officers

231. Sworn personnel in the SEPTA Police Department include the following ranks: patrol officer, corporal, sergeant, lieutenant, captain, deputy chief, and chief. These officers are commissioned to be police officers pursuant to Act 120.

232. Since 1991, SEPTA policy has required that incumbent sworn employees of all ranks in SEPTA's Transit Police Department take and pass a physical fitness test every six months. Despite this policy, there was evidence introduced at trial that incumbents are not always retested every six months.

233. The incumbent physical fitness testing program is based upon the same study relied on by SEPTA for its applicant physical fitness testing program. The components of SEPTA's physical fitness test for applicants that are being challenged in this case are identical to the components of SEPTA's physical fitness test that have been administered to incumbent SEPTA transit police officers since 1991.

234. The gym-based components of the physical fitness test administered to incumbents since 1991 are the same as the gym- based components of the physical fitness test administered to applicants since 1991.

235. The aerobic capacity component of the physical fitness test administered to incumbents since 1991 is conducted on a treadmill. According to SEPTA, the passing score on the treadmill test administered to incumbents measures the same level of aerobic capacity as the passing score on the 1.5 mile run administered to applicants.

236. Beginning in 1991, physical fitness tests for incumbent SEPTA transit police officers were administered at the Benjamin Franklin Clinic pursuant to a contract between SEPTA and that entity. The Benjamin Franklin Clinic closed in February 1997.

237. Since SEPTA has adopted the criterion tests for transit police officer applicants, incumbent officers who fail a component of the physical fitness test are given the option of taking a corresponding criterion test. The components of the criterion test offered to incumbents who fail a component of the physical fitness test are identical to the components of the criterion test administered to transit police officer applicants since 1996.

238. The corresponding criterion test for the aerobic capacity test on the treadmill is a run of 1.5 miles in 12 minutes or less. The criterion tests for the grip strength component of the gym-based test are the weapon fire and the dummy drag tests. The corresponding criterion tests for the push-up, pull-up, sit-up and bench press components of the gym-based test are the turnstile jump, barrier surmount and dummy drag tests.

239. SEPTA policy requires incumbent transit police officers who fail any component of the physical fitness test to be retested on the failed components within three months.

240. For each component of the physical fitness test that an incumbent transit police officer fails, an interim goal is set for that officer. The incumbent officers were provided with interim goals in order to allow these officers, who were not hired under SEPTA's rigorous physical fitness test, to gradually work toward and achieve the fitness standards that the applicants need to achieve — by 1996, 86% of the officers hired prior to the Davis test reached SEPTA's physical fitness standards. When the physical fitness tests for incumbent tran-

sit police officers were administered at the Benjamin Franklin Clinic, the interim goal was determined by negotiation between the incumbent transit police officer and staff at the Benjamin Franklin Clinic.

2. The Pass Rates of Incumbent Officers

241. All incumbent officers, regardless of rank, including SEPTA's Chief of Police, are required to pass the physical fitness test because any such officer is subject to being called out to perform patrol duties and must therefore be prepared to carry out these duties.

242. SEPTA's own internal memoranda document that incumbent transit police officers of all ranks have failed SEPTA's physical fitness test — the same physical fitness test administered to applicants that is at issue in this case.

243. One such document, dated September 22, 1995, indicates that between July 1, 1994 and August 22, 1995 the percentage of uniformed personnel who failed the fitness test was as follows: 10% of all officers between the ages of 20 to 30; 30% of all officers and 12% of all supervisors between the ages of 30 and 40; 45% of all officers and 52% of all supervisors between the ages of 40 and 50; and 55% of all officers and 40% of all supervisors between the ages of 50 and 60.

244. Other internal SEPTA documents establish that these incumbent officers and supervisors often failed the physical fitness test on more than one occasion during this time period.

245. According to a chart introduced by the plaintiffs at trial, (Pls.' Ex. 106), since SEPTA began administering its physical fitness test to incumbent transit police officers, the following percentages of such officers have failed the following components of the physical fitness test on at least one occasion:

Component	Percentage of Officers Who Failed
Any Component	69.97%
Aerobic Capacity	62.20%
Push-Up	41.64%
Pull-Up	29.79%
Sit-Up	29.35%
Bench Press	17.35%
Grip Strength	11.26%

These percentages, however, do not appear to be correct. The employee of the United States Justice Department who prepared this chart testified that, although the numbers on the chart purport to represent the percentage of officers who failed any component or a particular component of the physical fitness test at any time, the chart actually could have counted the same officer a number of times if this officer failed the test a number of times. In addition, the chart does not reflect whether an officer failing a test component passed the corresponding criterion task test. Officers considered to have passed by SEPTA may have been erroneously included in this chart by the plaintiffs. Thus, this evidence is not entitled to much weight.

246. According to another chart introduced into evidence by plaintiffs, (Pls.' Ex. 107), 182 such officers have failed the aerobic capacity component of the test (by scoring less than 42 mL/kg/min) on at least one occasion. This chart, however, does not indicate whether the failure rate of any component by the entire group on a percentage basis improved over time. Therefore, this chart cannot demonstrate the progression of incumbents with regard to physical fitness test results.

In addition, this chart incorrectly indicates that 182 officers failed the aerobic capacity component at least once. This number of 182 actually represents the number of test events on which there was a failure, not the numbers of officers who failed — one officer could have been counted ten times if the officer failed the test ten times. Thus, this chart is not entitled to much weight.

247. Plaintiffs' charts also fail to indicate whether the officer who failed a particular fitness component was hired before or after the implementation of the physical fitness testing for applicants. In addition, plaintiffs' charts fail to indicate whether the officer, who they considered to be failing, passed the interim goals that had been set by SEPTA management.

248. In contrast to the test results offered by plaintiffs, defendant introduced evidence which established that from

1991 to 1996, 96.2% of SEPTA officers have passed the grip strength component, 98.1% of SEPTA officers have passed the bench press component, 92.9% of SEPTA officers have passed the sit-up component, 85.0% have passed the pull-up component and 81.1% have passed the push-up component.

3. The Implementation and Administration of Incumbent Testing

249. When incumbent testing was first introduced, SEPTA would discipline incumbent officers for failing to meet their interim goals. However, the patrol officers' union objected to such discipline, claiming that the disciplinary component of SEPTA's physical fitness testing was never the subject of collective bargaining, and thus SEPTA could not unilaterally implement such testing. The union took SEPTA to arbitration over this matter and won. Thus, due to the opposition of the patrol officer's union, SEPTA was precluded from disciplining the patrol officers who failed the incumbent testing.

250. Because SEPTA was unable to discipline officers who failed incumbent fitness testing, Chief Evans attempted to gain compliance with the incumbent fitness standards by offering an incentive whereby officers would receive $50.00 each time they passed their interim fitness goals, with a maximum of $200.00 per year. SEPTA additionally offered to reimburse officers for gym memberships. This incentive program for incumbent officers was implemented with the union's concurrence.

251. Given that SEPTA does not have the ability to discipline its incumbents who fail to meet interim fitness goals set by SEPTA, Chief Evans believes that those few officers who repeatedly fail their incumbent testing do so because of a lack of effort, desire or motivation. Chief Evans has elected not to impose discipline on supervisors because he does not believe that half of the police department should be treated differently than the other half — the transit police officers who he cannot discipline.

252. Although SEPTA has never taken any steps to determine whether the incumbent officers who have failed the physical fitness test have adversely affected SEPTA's ability to

carry out its mission, Chief Evans testified that officers who are not passing their incumbent fitness examinations are not capable of performing all of their policing duties and that a lack of fitness and inability to meet fitness standards has resulted in on-the-job injuries. For example, Chief Evans testified to an incident where a SEPTA officer, who was not meeting her interim fitness goals, was thrown into the track area of a train station by an intoxicated individual. Chief Evans believes that her lack of fitness contributed to her being thrown onto the tracks.

4. The Effect of Incumbent Testing

253. Lt. Maslin, who is in charge of supervising patrol officers and is intimately familiar with the scores that particular officers have received on their physical fitness tests, has observed the impact of the physical fitness testing program for incumbents. In his estimation, the program has resulted in "higher caliber officers" who are more vigilant in patrol and who are better able to effectuate backups and assists to fellow officers.

254. Lt. Maslin has observed the progress of the incumbents in moving toward and meeting SEPTA's fitness standards because he is in charge of computerizing the fitness data for the incumbent officers. From this base of knowledge, Lt. Maslin was able to discern that officers arriving at calls who were meeting SEPTA's standards were in better shape than those officers arriving at the scene who were unable to meet the standards.

255. Since the implementation of this fitness program, Part I felony offenses, *i.e.*, homicide, rape, robbery, aggravated assault, burglary, theft and auto theft, are down by approximately 70%. Lt Maslin believes that the fitness program has contributed to this reduction in crime.

5. The Performance of Incumbents

256. SEPTA has promoted incumbent officers who have failed some or all of the components of the physical fitness test at any time. Since July 1994, the Chief of SEPTA Transit Police Department has had the authority to remove candi-

dates from promotional lists for failing to achieve their interim fitness goals. Despite the authority to remove officers from the promotional lists, no SEPTA officer has ever been removed from a promotional list for failure to pass physical fitness testing for incumbents. Nevertheless, only ten officers who have failed their physical fitness tests have ever been promoted.

257. SEPTA has also given special recognition to incumbent officers who have failed the physical fitness test, such as Officer of the Quarter. SEPTA has also awarded numerous commendations for outstanding service to officers who have failed at some point in time any component of their physical fitness testing.

258. SEPTA has also given satisfactory performance evaluations to incumbent officers who have failed one or more components of the physical fitness tests. However, these performance evaluations were only completed for supervisory police personnel, *i.e.*, sworn employees above the rank of transit police officer. Moreover, these evaluations were not specially created for the Transit Police Department, rather these evaluations were used for general supervisory, administrative and management employees throughout the SEPTA system.

259. SEPTA has also never disciplined or sought to discipline, terminated, removed, reassigned, suspended from duty or demoted any transit officer for failing to perform the physical requirements of the job.

I. Selection of Applicants who Failed the Physical Fitness Test

260. SEPTA has selected two applicants who failed the physical fitness test.

261. For example, Officer Thomas was hired in 1991 despite the fact that she did not complete the 1.5 mile run in 12 minutes and failed the bench press, sit-up and push-up components of SEPTA's physical fitness test for applicants. Officer Thomas has gone on to become a decorated officer who has repeatedly been nominated for awards such as Officer of the Year and Officer of the Quarter. In fact, SEPTA has commended Officer Thomas for her outstanding perfor-

mance as a police officer. Moreover, Officer Thomas serves as one of SEPTA's two defensive tactics instructors.

262. SEPTA also hired Officer Baxter in 1991 despite the fact that she failed the bench press and push-up components of SEPTA's physical fitness test for applicants.

263. At the time these two individuals were hired in 1991, the Human Resources Department of SEPTA administered the applicant test; the SEPTA Transit Police Department was not involved in the administration of the 1991 test. Thus, if Officers Thomas and Baxter were hired without successfully passing all components of the physical fitness test, the error occurred outside the control of the SEPTA Transit Police Department.

J. The Statistical Analyses Conducted By Drs. Griffin and Siskin Demonstrating the Job-Relatedness and Business Necessity of SEPTA's Physical Fitness Test

264. After this litigation commenced, SEPTA retained statisticians, Bernard Siskin, Ph.D., and David Griffin, Ph.D., to submit expert reports which examine the statistical relationship between the components of SEPTA's physical fitness test on the one hand and the number of arrests and "arrest rates" on the other.[5]

265. In addition, Drs. Griffin and Siskin conducted a "commendation analysis" which demonstrates the relationship between officers receiving commendations for outstanding acts in the performance of their duties and the aerobic capacity of these officers.

266. Drs. Griffin and Siskin also conducted a "perpetrator analysis" which calculates the estimated aerobic capacities of persons arrested by SEPTA officers for Part I crimes between 1991 and 1996 and compares those aerobic capacities with the aerobic capacities of SEPTA transit police officers.

267. For the following reasons, the Court finds the statistical analyses conducted by Drs. Griffin and Siskin establish

---

5. Dr. Siskin testified at trial as to the results of the studies and reports and the opinions expressed therein. Dr. Griffin only testified as to some of the underlying data.

that SEPTA's aerobic capacity requirement is job-related and consistent with business necessity.[6]

---

6. Drs. Siskin's and Griffin's analysis of arrests, arrest rates and commendations and their relationship to aerobic capacity is offered by SEPTA as evidence of validity under a criterion-related validation strategy. Evidence of the validity of a test or other selection procedure by a criterion-related validity study consists of empirical data demonstrating that the selection procedure is predictive of or significantly correlated with important elements of job performance. The hallmark of criterion-related validity is empirical data establishing a statistically significant correlation between performance on the test and objective measure or "criteria" of job performance. Under a criterion-related validation strategy, a proponent must show two elements of correlation. *See Dickerson v. United States Steel Corp.*, 472 F. Supp. 1304, 1349 (E.D. Pa. 1978) ("In addition to requiring that the correlations of the test battery to the criteria be statistically significant, [the] guidelines require that the correlations indicate practical significance") (citation omitted). The first is practical significance, which is the degree to which the test scores relate to job performance, and is usually measured by a "correlation coefficient." The second is statistical significance, which is the measure of confidence that can be placed on the practical significance. In other words, the statistical significance expresses the probability that a particular correlation coefficient occurred by chance. *See Hamer v. City of Atlanta*, 872 F.2d 1521, 1525-26 (11th Cir. 1989). In *Ensley Branch of NAACP v. Seibels*, 616 F.2d 812 (5th Cir. 1980), the former Fifth Circuit explained a few of the "statistical concepts" that underlie a criterion-related study. Because the Court believes that such an explanation would be helpful here, that portion of the *Seibels* opinion will be repeated here:

> Statistically, the degree of correlation between two variables (*e.g.*, entrance exam scores and subsequent school grades) is expressed as a "correlation coefficient" on a scale running from +1.0 to -1.0. A perfect positive correlation (*e.g.*, entrance exam scores exactly predict subsequent school grades, with the higher exam scores predicting the best grades) would be expressed as +1.0 and a perfect negative correlation (*e.g.*, entrance exam scores exactly predict subsequent school grades, except in reverse, with the lower exam scores predicting the best grades) would be expressed as -1.0. Where the two variables had absolutely no relationship to each other, the correlation coefficient would be .0. The closer a correlation coefficient is to either +1.0 or -1.0, the "higher the magnitude" of the correlation; and the closer it is to .0, the "lower the magnitude." Mueller, Schuessler & Castner, *Statistical Reasoning in Sociology*, 2d Ed., at p. 315. Because a purely random drawing of a sample is liable to produce a correlation coefficient which is somewhat off an absolute .0, the concept of statistical significance is relevant. The concept is tied to the statistical theory of probability and is dependant upon the number of the people in the

268. Dr. Siskin's analysis found that officers with an aerobic capacity of 42 mL/kg/min[7] or higher had statistically significant higher numbers and rates of arrests with respect to Part I crimes and all offenses than officers who were below SEPTA's aerobic capacity requirement of 42 mL/kg/min. In sum, officers who met or exceeded SEPTA's aerobic capacity requirement made more arrests, particularly Part I arrests, than those officers who had an aerobic capacity below SEPTA's requirement of 42 mL/kg/min and were more likely to make an arrest per incident, especially for Part I crimes, than those officers below 42 mL/kg/min.

269. Dr. Siskin found that the relationship between aerobic capacity and arrests and arrest rates was linear. This means that the higher the aerobic capacity of the officer, the higher you would predict their number of arrests and their arrest rate. This demonstrated linear relationship was established both for all offenses and especially for the more serious Part I offenses.[8] These findings were statistically significant at less than .05 and in many cases less than .001, thus meeting the significance requirement of .05 of the *Uniform Guidelines*.

270. Drs. Siskin and Griffin calculated correlation coefficients on three different bases: a test event basis; an officer basis; and an officer average basis. The test event basis looks at each discrete physical test event; the officer basis looks at

---

sample. Generally, if a correlation coefficient is so low that, on the basis of the random sample size involved, more than 1 in 20 random drawings could be expected to produce a correlation at least as great, that correlation coefficient is considered not to be statistically significant, or simply to be the same as a correlation coefficient of .0. On the other hand, if the obtained coefficient could be expected to reoccur no more than once in 20 random drawings, it is considered statistically significant, the statistical indication for which p could not be expected to occur by chance more than once in 100 random drawings is expressed as p Mueller, et al. pp. 394, *et seq. Seibels*, 616 F.2d at 817 n.13.

7. Although SEPTA's standard is 42.5 mL/kg/min, the Court will simply refer to it as 42 mL/kg/min throughout this section of the opinion because Dr. Siskin referred to the standard as 42 mL/kg/min.

8. As part of their studies, Drs. Siskin and Griffin analyzed the statistical relationship between aerobic capacity and Part I arrests and overall arrests. Overall arrests included Part I arrests.

an officer's average performance; and the officer average basis looks at the average performance of a group of officers. Under the officer average basis, Dr. Siskin viewed the data by grouping officers at various aerobic capacity levels.

271. Calculating the correlation coefficients on the officer average basis, aerobic capacity was highly predictive of the average number of arrests and arrest rates of all officers at that aerobic capacity level for all offenses and Part I offenses. Dr. Siskin found the correlation coefficient between aerobic capacity and the average arrest rate of officers to be approximately 0.4 and the arrest rate for the more serious Part I offenses to be .52. The data demonstrated that one can reasonably expect that, on average, officers with a higher aerobic capacity will convert more arrest opportunities into arrests, and make more arrests, both for Part I offenses and for all offenses, than officers with a lower aerobic capacity.

272. Although Dr. Siskin admitted that "traditional validation is done at an individual level of analysis, that is with data collected from individuals and interpreted as predictions of individual criterion performance," Dr. Siskin expressed his professional opinion that the officer average basis has utility in this case because it helps express the practical implications of the studies that he and Dr. Griffin conducted. In other words, the officer average basis helps demonstrate how arrests over the period of time from 1991 through 1996 would have increased at SEPTA if the officers, who had an actual aerobic capacity below 42 mL/kg/min during this period, had an aerobic capacity at or above 42 mL/kg/min during this same period.

273. On a test event basis, the highest reported correlation between passing SEPTA's test and any of SEPTA's criterion measures for patrol officers is .131 (the correlations between passing all components of SEPTA's test and arrests per year for Part I crimes). The highest correlation between passing the aerobic capacity component of the test and any of SEPTA's criterion measures for patrol officers on a test event basis is .107.

274. On an officer basis, Dr. Siskin recorded correlation coefficients as high as .22. He also testified that this correlation

was uncorrected and that psychometricians normally would correct such a correlation coefficient for "restriction of range" and "criterion unreliability." If these corrections had been done here, the .22 correlation would increase to approximately to .33.

275. Dr. Siskin also found that the likelihood of receiving a commendation for "street" patrol officer performance was statistically significantly higher if the officer's aerobic capacity met or exceeded 42 mL/kg/min. Dr. Siskin reviewed 207 commendations that were awarded for the period of 1994 through 1996 and found that 96% of the commendations went to officers who had an aerobic capacity greater than 42 mL/kg/min; these officers had an average aerobic capacity of 46 mL/kg/min. Furthermore, 198 of the commendations studied involved an arrest, with 116 having an explicit reference in the commendation document to a foot pursuit, use of force or other physical exertion.

276. Dr. Siskin's testimony also showed, when comparing officers who were always at 42 mL/kg/min or over to officers who were always under 42 mL/kg/min, the higher aerobic capacity group had a 57.1% "arrest rate" advantage in the more serious Part I crimes and 28% greater arrest rate for all offenses. Dr. Siskin also pointed out that the data showed that officers always at 42 mL/kg/min or above made three times (151%) the actual number of Part I arrests and 75% more actual overall arrests when compared to officers who never met the 42 mL/kg/min requirement.

277. During the course of the trial, the plaintiffs, primarily through the testimony of Dr. Zedeck, attempted to undermine the validity of Drs. Siskin's and Griffin's studies by pointing out alleged flaws in the studies. Dr. Siskin, however, demonstrated that such flaws did not actually exist and that if these flaws did exist, the flaws did not undermine the validity of the studies.

278. Dr. Siskin addressed the plaintiffs' concerns about "contaminating factors" — factors which could have upset the statistical relationships discovered by Drs. Siskin and Griffin including age, tenure and learning by controlling for rank and assignment. Dr. Siskin did this through a "regression analysis"

that adjusted the studies for zone, shift and rank (patrol officers versus sergeants). Regression analyses allow a statistician to compare people who are similarly situated with respect to their assignments.

279. The regression analysis conducted by Dr. Siskin showed that the differences between the officers who achieved 42 mL/kg/min or higher versus the officers who never met 42 mL/kg/min was still statistically significant in the number of Part I arrests made and the arrest rate for Part I crimes and the arrest rates for all crimes. Specifically, after the regression analysis was run, Dr. Siskin's data showed a 14% advantage in the overall arrest rate for officers at or above 42 mL/kg/min, a 32% arrest rate advantage for officers at or above 42 mL/kg/min for Part I crimes, as well as a significant difference in the number of Part I arrests made by officers meeting or exceeding SEPTA's aerobic capacity standard.

280. Dr. Siskin also testified that rotating officers within various zones and tours and through different beats would have no effect on his conclusions because beat assignments are not correlated to an officer's aerobic capacity.

281. Dr. Siskin, in performing his studies, controlled for special units and unfounded incidents and found that these variables, like his other controls, did not effect the outcome of his studies.

282. Dr. Siskin testified that beat assignments can be considered "random noise" that would only obscure and lower the observed correlation coefficients and statistical significance. Notwithstanding this "noise," all of Dr. Siskin's studies were significant at either less than the .05 level or less than the .01 level, and in many instances less than the .001 level. Dr. Siskin testified that running a partial correlation for "beat" assignments would have only raised the correlation and the level of statistical significance.

283. Dr. Siskin pointed out that random errors in measurement or errors in the data can be considered the same as random noise. Dr. Siskin testified that there was no reason to believe that these types of errors — errors in measurement, data, attribution, *etc.* — will favor either a high aerobic capac-

ity group or low aerobic capacity group, hence they are random with respect to aerobic capacity and act as random noise.

284. Dr. Siskin explained that once a statistically significant relationship is found, random noise only acts to suppress the correlations between aerobic capacity and the criterion measures. In essence, random noise or random errors do not create a relationship, rather this randomness only masks such a relationship. Indeed, Dr. Siskin testified that once a correlation is found and adjustments are made for random noise or error, the statistical corrections will raise the correlation. Consequently, in this case, Dr. Siskin found that the observed correlations were an underestimation of the true relationship between meeting SEPTA's aerobic capacity requirement and making Part I arrests, overall arrests and arrest rates.

285. Dr. Siskin testified that while corrections for random noise would "clearly increase the correlations" so that the estimates of the correlations that he obtained in this case were actually too low, he did not make these corrections because the best measure of practical significance is found through regression analysis and expectancy tables, which estimate the effect of meeting SEPTA's aerobic capacity standard of 42 mL/kg/min relative to not meeting this standard. More significantly, these estimates are unaffected by random noise.

286. For example, Dr. Siskin pointed out that the 5.9% arrest rate advantage found in his regression study, which will be discussed below, would remain the same even if the correlation coefficients were corrected.

287. Dr. Siskin was asked whether or not any of the plaintiffs' criticism concerning measurement or methodology would affect his conclusions. Dr. Siskin noted that if he did not find a relationship between aerobic capacity in Part I arrests or overall arrest rate, then he might have been concerned. His concern, however, would have been that flaws in measurement or methodology would have obscured the relationship, and any conclusion that there was not a relationship between aerobic capacity and arrests might have been a mistaken conclusion. However, the fact that the data clearly and consistently showed a statistically significant relationship between meeting SEPTA's aerobic capacity standard of 42 mL/kg/min and

arrests and arrest rates, even when controlling for assignments, demonstrated that the conclusions were very solid.

288. Consequently, the import of Dr. Siskin's testimony was that once a relationship between aerobic capacity and arrest and arrest rates was found in the data, any controls for random noise, measurement errors or any other factors random with respect to an officer's aerobic capacity levels would only have raised the correlation and increased the statistical significance which was already at less than .05 and less than .01 levels.

289. Specifically, Dr. Siskin addressed the Court's concern that perhaps an officer could avoid using physical exertion in making an arrest or, appropriately, opt not to make the arrest. Dr. Siskin explained that his study did not simply look at physical arrests but at total arrests, thus the first scenario could not affect his results.

290. Dr. Siskin pointed out that the issue of judgment as to when to make an arrest was not a concern for his study because the results were essentially being driven by Part I arrests and Part I arrest rates, and it was hard for him to conceive that a SEPTA officer was not supposed to make an arrest in a robbery, rape, assault or theft circumstance — the types of serious crimes that are reflected in actual Part I arrests and Part I arrest rates.

291. With respect to arrests other than Part I arrests (*i.e.*, for offenses that could include prostitution, vagrancy, public urination and things of that nature), Dr. Siskin pointed out that there was no evidence that officers with low aerobic capacity would not make the arrest, and that officers with a high aerobic capacity would make the arrest for prostitution, vagrancy, or urination. Dr. Siskin testified that there was no data to show that with respect to high aerobic capacity individuals versus low aerobic capacity individuals, there is a judgment factor that falls in favor of either aerobic capacity group.

292. Dr. Siskin further testified that the Court's concern the ability to avoid a physical confrontation in making an arrest or judgment of when to make an arrest — would not affect his study because Part I crimes and Part I arrest rates were driving the results of his studies.

293. Dr. Siskin also addressed plaintiffs' concerns that the studies should have focused solely on physical arrests. Dr. Siskin explained that focusing on just physical arrests would have been biased in favor of SEPTA. Dr. Siskin noted that a study could be conducted which would focus in on arrests which could require physical exertion. He testified, however, that if one would focus on arrests that require physical exertion, the results would have been to raise the correlations and statistical significance he found.

294. Dr. Siskin also addressed the plaintiffs' concern that he did not control for rank, *i.e.*, his initial study included both sergeants and patrol officers, since sergeants are out in the transit system making arrests. Nevertheless, Dr. Siskin addressed this concern and testified that the inclusion of sergeants did not affect his results. He controlled for rank in two ways. One method was through a regression analysis in which he controlled for whether the officer was a sergeant or a patrolman. Further, Dr. Siskin pointed out that he ran all studies looking only at patrolmen and none of the findings changed. Specifically, Dr. Siskin testified that his results — officers meeting or exceeding SEPTA's aerobic capacity standard outperformed officers who failed to meet SEPTA's aerobic capacity standard — were not being driven by the inclusion of sergeants. Dr. Siskin found the same statistical relationship by simply looking at patrol officers. In sum, Dr. Siskin stated that the theory that sergeants were somehow different and were possibly driving the results was simply not accurate.

295. Dr. Siskin stated that the truest measure of estimating the effect of aerobic capacity on arrests and arrest rates was to look at the officer's field performance within time bands closely proximate to the test of aerobic capacity rather than averaging the officer's aerobic capacity over the course of his career. This method was described as the test event basis and was criticized by Dr. Zedeck.

296. Dr. Siskin explained that too much information is lost by averaging the officer's aerobic capacity over the course of his career. Thus, Dr. Siskin's approach was geared to measure the effect of aerobic capacity at the time an aerobic

capacity test was taken and estimate its effect on the officer's field performance at roughly that particular time. Dr. Zedeck's approach would be to average all of the officer's aerobic capacity tests over a 6-year period and then determine the aerobic capacity effect on arrest rates and overall arrests.

297. Dr. Siskin demonstrated through the use of Dr. Zedeck's tables, (noted as Exhibit "A" to Dr. Zedeck's rebuttal report), how much useful information about the relationship of aerobic capacity to arrests, arrest rates and Part I arrests is lost through this type of averaging. Exhibit "A" was a table that concerned commendations but nonetheless demonstrated that Dr. Zedeck's approach would effectively conceal the upward changes in aerobic capacity that ultimately led to the commendation events. For example, Dr. Siskin pointed out that for Officer Felix Adorno, his aerobic capacity varied and progressively included 39, 41, 47, 45, 47, and 44 mL/kg/min, yet at the time he received his commendation he was at 47 mL/kg/min. Averaging Felix Adorno's aerobic capacity would conceal the changes in his aerobic capacity, and thus obscure the effect of aerobic capacity on Officer Adorno's field performance.

298. Dr. Siskin testified that the test event basis, *i.e.*, measuring the effect of aerobic capacity and its relationship to field performance at the time an aerobic capacity test was given to an officer, was the best estimate of how aerobic capacity related to the various arrest parameters. Dr. Siskin testified that the test event basis would actually lower the correlations compared to Dr. Zedeck's proposed officer basis because a person has to try to predict a single event at a single point in time under the test event basis rather than the average performance of an officer as with the officer basis. However, the test event basis is the most accurate way of measuring the effect of aerobic capacity on the field measures of overall arrests, arrest rates and Part I arrests. An officer basis analysis would yield a statistically biased — too low — estimate of the relationship between aerobic capacity and arrests.

299. Because the test event basis may have some inter-officer correlation, Dr. Siskin testified that while the estimate of the effect is accurate, the test of significance is not perfectly

accurate. Hence, he testified that he conducted additional tests to assure that the inter-officer correlation was not creating the statistical significance. Dr. Siskin conducted several tests that confirmed that the statistical significance that he discovered on the test event basis was always real.

300. Dr. Siskin expressed complete confidence that the true statistical significance of the relationship between aerobic capacity and Part I arrests, overall arrests and Part I arrest rates were significantly below the .05 level that is recommended by the *Uniform Guidelines*.

301. Dr. Siskin testified that in this case, correlation coefficients are not the proper focus in determining practical significance. Instead of using a correlation coefficient to determine the practical significance, Dr. Siskin testified that the appropriate measure of practical significance is the estimated impact of the effect of aerobic capacity on Part I arrests, overall arrests and arrest rates. In this regard, Dr. Siskin found that SEPTA could expect a half percent increase in Part I arrests for every increase in mL/kg/min of aerobic capacity and that such an effect was linear.

302. Dr. Siskin testified that the correlation coefficient issue was in some sense a "red herring" because the important question was the practical significance, *i.e.*, the predicted increase in arrests for the officers who did not meet SEPTA's standard if they performed like those officers who maintained an aerobic capacity of 42 mL/kg/min or above. Dr. Siskin explained that the best indicator of the practical significance of the relationship between arrests, arrest rates and aerobic capacity is demonstrated through regression analysis which explicitly measures the expected gain, rather than looking at the level of the correlation coefficient in which the value changes depending on what is predicted (an officer's performance at a point in time, an officer's performance over time or the performance of a group of officers over time) or whether you correct the correlation upwards to correct for restriction in range and criterion unreliability.

303. Under his regression analysis, Dr. Siskin demonstrated that for the period of 1991 through 1996, SEPTA could have achieved 470 additional arrests — 70 of which were Part

I arrests for serious crimes — if the aerobic capacity of all the officers was 42 mL/kg/min or above for this time period. These findings reflect a 10% increase in Part I arrests and a 4% increase in the overall arrest rate. This analysis was based on a regression analysis that took into account all relevant variables, including rank, zone and tour and assignments to special units. Dr. Siskin testified that taking these variables into account, the statistical relationship and predictive nature of aerobic capacity remained significant and demonstrates that meeting SEPTA's aerobic capacity standard of 42 mL/kg/min consistently predicted higher arrests and arrest rates for Part I offenses.

304. Dr. Siskin stated that it is well known and can be proven mathematically that if you are measuring the utility of tests, correlation coefficients are an inappropriate measure.

305. Dr. Siskin's regression study is completely in accord with the SIOP Principles. The SIOP Principles specifically state that the "slope of the regression line" and "expectancy tables" are acceptable and may be preferable to correlation coefficients in determining the usefulness of a test:

> [When multivariate techniques are used, the number of cases should be large relative to the number of variables. The analysis should provide information about the strength of the relationship, usually a coefficient of correlation. *Other methods (such as the slope of the regression line, expectancy tables, or the percentage of misclassifications) are acceptable and may be preferable in many situations. The analysis should also give information about the nature of the relationship and how it might be used in prediction.*

SIOP Principles at 15 (emphasis added).

306. Defendant's Demonstrative Exhibit 12, "Regression Adjusted Predicted Arrest Increase for Officers Below 42+ ml", is a graphic depiction of the total arrest increase that was predicted by the regression analysis - 469 overall arrests. Defendant's Demonstrative Exhibit 12 depicts the slope (.039) of the regression relationship which is statistically significant at less than .001.

307. Dr. Siskin testified that in view of the linear relationship between aerobic capacity and the arrest parameters any cutoff score can be justified since higher aerobic capacity levels will get you more field performance.

308. From a statistical perspective, the data supports any cutoff score because in a linear relationship, an increase in one variable is accompanied by an increase in the other variable (*i.e.*, more is better), and therefore you are entitled to choose how much more you desire.

309. Dr. Siskin also described the commendation study that he conducted. Dr. Siskin reviewed 207 commendations and found that 96% of the officers receiving commendations had an aerobic capacity level of 42 mL/kg/min or greater. The mean aerobic capacity for officers receiving the commendations was 47 mL/kg/min. Dr. Siskin's analysis revealed that the receipt of a commendation was more likely to be associated with a higher aerobic capacity than a lower aerobic capacity.

310. Dr. Siskin pointed out that of the 207 commendations, 116 were clearly coded as having some indication of a pursuit, use of force or other physical exertion. These are identified as "Physicality" related commendations. However, a review of defendant's summary of the 207 commendations shows that 96% of the commendations involved an arrest. The use of the word "Physicality" only refers to the description that was contained in the underlying commendation document. Consequently, the column in defendant's Exhibit 52(b) that indicated "No Physicality" did not mean that the commendation was given for activities other than apprehensions and arrests. In fact, a review of the defendant's Commendation Summary shows that only six commendations were given for patrol officer work other than arrests, apprehensions, disarming suspects, use of force, foot pursuit or some other officer duty requiring physical exertion. Clearly, the commendations that Dr. Siskin studied were given for outstanding transit patrol officer work in the area of arrests and apprehensions, since 96% of the commendations involved an arrest, regardless of how they were coded in defendant's summary.

311. In the rebuttal report of Drs. Siskin and Griffin, Dr. Griffin undertook a review of the actual commendations and

concluded that the Commendation Summary was accurate and faithful in its description of the arrest event that led up to the commendation.

312. Dr. Siskin testified that he did a statistical test to determine whether the award of a commendation was statistically associated with aerobic capacity of 42 mL/kg/min or higher. Dr. Siskin found a statistically significant relationship in that an officer was less likely to receive a commendation if the officer had a lower aerobic capacity (less than 42 mL/kg/min) than if the officer maintained a higher aerobic capacity (42 mL/kg/min or greater).

313. In connection with his commendation study, Dr. Siskin conducted a statistical comparison of the aerobic capacity distribution of the officer work force and compared it to the aerobic capacity of the commended officers. The mean aerobic capacity (47 mL/kg/min) for the commended officers when compared to the entire officer population (44 mL/kg/min) was statistically significantly higher at the .01 level.

314. Dr. Siskin also studied 953 perpetrators who had been arrested for committing Part I crimes in order to determine their aerobic capacity. The analysis was based upon the sex, race and age of the perpetrators. Dr. Siskin utilized a study (the "Vogel Study") provided by one of defendant's experts, Dr. Moffatt, in order to develop a statistical prediction of the aerobic capacity levels of the 953 perpetrators who were apprehended during the years 1991-1996. Based on his analysis, Dr. Siskin was able to provide an estimate of the aerobic capacity of the 953 perpetrators who were caught or apprehended. The mean age of the arrested perpetrators was 26.3 yrs.

315. Dr. Siskin's analysis showed that 51.9% of the perpetrators were estimated to have an aerobic capacity of 48 mL/kg/min, and only 27% of the perpetrators were estimated at or below 42 mL/kg/min.

316. Dr. Siskin also conducted a study of the aerobic capacity of the SEPTA officers that apprehended perpetrators of Part I crimes in the SEPTA transit system.

317. This analysis can be found at defendant's Exhibit 53(d). Dr. Siskin studied 382 Part I arrests for the period of 1994-1996. Dr. Siskin found that the arresting SEPTA transit police officers maintained a mean aerobic capacity of 46.8 mL/kg/min; whereas, the aerobic capacity of the SEPTA transit patrol officer population was approximately 43.9 mL/kg/min. The aerobic capacity of the SEPTA transit police officers who apprehended the Part I criminals during the years of 1994 through 1996 was found to be statistically significantly higher (at the 0.01 level) than the general SEPTA patrol officer population. Furthermore, 94% of the arresting patrol officers in this study maintained an aerobic capacity that exceeded 42 mL/kg/min. Only SEPTA patrol officers who made arrests were studied. Therefore, of 382 possible matches between a perpetrator and an arresting officer, there were 281 cases of SEPTA transit patrol officers making the arrests.

318. Dr. Siskin stated that the outcomes of his perpetrator studies were neither surprising nor unexpected since the data showed a consistent pattern indicating that the arrest rates and actual arrests were higher for officers who maintained 42 mL/kg/min or greater, and thus Dr. Siskin would expect that the officers making the arrests would have higher aerobic capacities than the general SEPTA transit officer population.

319. Dr. Siskin was also asked to conduct an analysis of SEPTA's muscular strength and endurance tests which were known as the gym-based components of SEPTA's physical abilities test. His analysis used the same methodology — test event basis — that was described in assessing the relationship between aerobic capacity and overall arrests, Part I arrests and arrest rates. Dr. Siskin's findings were summarized in defendant's Exhibits 53-E and 53-F. Defendant's Exhibit 53-E was Dr. Siskin's initial study which looked at the relationship between arrests and passing individually the bench press, pull-up, sit-up, grip strength and the entire battery of muscular strength and endurance tests. Dr. Siskin's study found a statistically significant relationship between passing the various gym-based components and making Part I arrests and arrest rates. Further, defendant's Exhibit 53-E demonstrates that

passing the battery of muscular strength endurance tests and maintaining an aerobic capacity of 42 mL/kg/min or greater was statistically significantly related to the actual number of arrests for all crimes, Part I crimes and the arrest rates for all crimes and Part I crimes. The significance levels were either less than 0.05 or less than 0.01, as more fully described in 53-E. Again, the patterns were similar to those that were found when looking at the relationship between maintaining 42 mL/kg/min of aerobic capacity and the various criterion measures.

320. Dr. Siskin also did a regression study with respect to the gym-based components, controlling for tour, zone and rank, to determine whether or not the muscular strength and endurance tests still had a statistically significant relationship to any of the arrest parameters that he was studying. The regression analysis is described in defendant's Exhibit 53-F and showed that again there was a statistically significant relationship to making Part I arrests for those officers who met all the gym-based standards and who maintained an aerobic capacity of 42 mL/kg/min or higher. These officers made more Part I arrests than those officers who failed the gym-based tests.

321. Dr. Siskin was also asked to analyze from a statistical perspective Dr. McArdle's proposal that "relative fitness"[9] as opposed to absolute aerobic capacity would predict arrests or arrest rates.

322. For example, Dr. Siskin noted that based upon Dr. McArdle's model, a female at 36 mL/kg/min is considered as fit as a male who is at 42 mL/kg/min because the female would be at the fiftieth percentile for all women and the male would be at the fiftieth percentile for all men. Dr. Siskin conducted a

---

9. In proposing an alternative test, Dr. McArdle suggested that SEPTA could test women and men based on relative fitness, that is, men and women would be considered to have the same fitness levels if their aerobic capacity scores placed them at the fiftieth percentile for women and the fiftieth percentile for men respectively despite the fact that their absolute aerobic capacity scores would be different — men at the fiftieth percentile for all men would have greater absolute aerobic capacity scores than women at the fiftieth percentile for all women.

series of regressions to determine whether relative fitness, rather than absolute aerobic capacity, was a variable that predicted or correlated with the field performance parameters that he was studying. Based on these regression studies, Dr. Siskin found that there was no statistical support whatsoever for the proposition that relative fitness correlated with or predicted field performance. In fact, the regression studies showed a negative gender effect, and thus these relative fitness standards were not predictive of performance whatsoever under these circumstances.

323. Dr. Siskin testified that defendant's Demonstrative Exhibits 14 and 15 showed the results of his study of Dr. McArdle's premise that relative fitness would predict performance in the various arrest parameters that he was studying. For example, defendant's Demonstrative Exhibit 14 shows that the arrest rate for males at 42 mL/kg/min of aerobic capacity was 23% and the arrest rate for females at 36 mL/kg/min was 7.7%, demonstrating that relative fitness does not predict field performance for SEPTA transit police officers. In addition, a review of Demonstrative Exhibit 14 shows that females in the 36 mL/kg/min to 41 mL/kg/min range, who under the Cooper standards would be expected to be at a "higher" level of fitness than a male of the same age category at 42 mL/kg/min, only attained a 9.8% arrest rate — a rate that is far below that of the arrest rate of males with 42 mL/kg/min of aerobic capacity. Dr. Siskin testified that there is nothing in the data that would support an argument that one should be looking at relative fitness as opposed to absolute values for aerobic capacity.

324. As was noted in defendant's Demonstrative Exhibit 14, Dr. Siskin held age constant for males and females at approximately 32 years. Thus, his findings completely contradict Dr. McArdle's assertions that relative fitness is a useful predictor. In fact, as Dr. Siskin testified with respect to defendant's Demonstrative Exhibit 15, there was a negative gender effect when one used the Cooper relative fitness model.

325. Dr. Siskin concluded his testimony by describing defendant's Demonstrative Exhibit 24 which showed that: (1) 100% of the officers who received the "Officer of the

Quarter/Year" award were at or above 42 mL/kg/min with a typical aerobic capacity of 45.1 mL/kg/min; (2) 96% of the officers who received commendations had an aerobic capacity in excess of 42 mL/kg/min and typically maintained an aerobic capacity of 46.6 mL/kg/min; (3) 75% of the individuals promoted to sergeant or lieutenant maintained an aerobic capacity of 42 mL/kg/min or greater, with a typical aerobic capacity of 43.3 mL/kg/min; (4) 76% of the perpetrators of Part I crimes who were arrested had an aerobic capacity of 42 mL/kg/min or higher with a typical aerobic capacity of 47.8 mL/kg/min; and (5) 94% of the officers who arrested Part I perpetrators, in the group he studied, had an aerobic capacity of 42 mL/kg/min or greater with a typical aerobic capacity of 46.8 mL/kg/min.

K. The Study of Dr. Robert Moffatt, Ph.D., Offered to Demonstrate the Job-Relatedness and Business Necessity of SEPTA's Physical Fitness Test

326. Subsequent to the filing of the *Lanning* administrative charges with the PHRC and the EEOC, SEPTA retained Robert Moffatt, Ph.D., an exercise physiologist, to defend SEPTA's physical fitness test.

327. Dr. Moffatt's study shows that there is more of a decrease in performance of certain gross motor skills after a period of anaerobic exercise for persons with lower aerobic capacities.

328. After accepting the assignment with SEPTA in March 1996, Dr. Moffatt visited Philadelphia in May 1996 and conducted interviews with SEPTA police officers. Dr. Moffatt questioned the officers about their job duties and obtained a tour of SEPTA's transit system to understand the nature and environment within which the police officers worked. In questioning the transit officers about their jobs, Dr. Moffatt discovered job duties that enabled him to perform a test that would demonstrate the predictive nature of the SEPTA aerobic capacity test.

329. During his two tours of the SEPTA system, Dr. Moffatt observed dramatic differences between the job duties of a SEPTA officer and those of other law enforcement offic-

ers with whom he had worked — the Citrus County, Florida Sheriff's Office and the Metropolitan Dade County, Florida Sheriff's Office. Dr. Moffatt noted that the SEPTA transit police force is predominately on foot patrol and arrives at various locations on foot. The SEPTA officers patrol alone and traverse a wide number of steps during their shifts.

330. In interviews with the SEPTA officers, Dr. Moffatt was told that one of the critical tasks of a SEPTA officer is running from one station to the next for officer assist calls. The officers also told Dr. Moffatt that they had to be prepared to fight or subdue a perpetrator upon arrival. Because this scenario was deemed a critical task, Dr. Moffatt decided to test for the amount of aerobic capacity that would be necessary to successfully engage in this task.

331. Dr. Moffatt wanted to determine through a simulation of a typical SEPTA backup/assist call how long it would take the officers to run from point A to point B. Protocols were devised for the testing of SEPTA transit police officers from which Dr. Moffatt could establish a pace for use in laboratory testing.

332. The protocols for the simulated runs were sent by Dr. Moffatt to SEPTA Captain Steven Harold on June 10, 1996. Officers were requested to take part in two different scenarios. From having spoken to the SEPTA officers, Dr. Moffatt was informed that their average officer backup or assist calls generally last from three to four minutes in duration. Therefore, Dr. Moffatt chose this time interval for his simulated test and had Captain Harold choose a "real-to-life" course that SEPTA officers routinely run. Thus, a concourse run was developed from the City Hall area to 11th Street. Each of the officers participating in the simulation ran two scenarios — an officer backup and an officer assist. The officer backup was an example of crowd control, and the officer assist was to aid an officer with the anticipation that a struggle might ensue upon arrival.

333. Although Dr. Moffatt learned that SEPTA officers often run distances of six to ten blocks, he told Captain Harold to be conservative and pick a "more typical" running scenario.

334. In order to ensure that the outcome would be random, the officers who ran the officer backup and officer assist calls did so in a mixed fashion, such that some officers ran the assist scenario first and other officers ran the backup scenario first. Each officer was given a rest period of approximately 60 minutes between the running of the backup and assist scenarios.

335. Captain Harold included a dummy drag at the conclusion of the first running simulation which occurred on June 25, 1996. Dr. Moffatt concluded that Captain Harold's inclusion of a dummy drag was insightful because it provided further proof of the decrement in work ability of SEPTA officers at the conclusion of running from location A to location B.

336. In order to obtain a "baseline" from the running group that participated in the first simulation on June 25, 1996, those individuals were brought back to engage in a dummy drag in a rested state so that a contrast could be drawn between how long it took them to do a dummy drag both before and after the running simulation. The second group of officers that participated in running simulations on September 5, 1996 completed a dummy drag before their runs to establish a baseline and then completed a second dummy drag at the conclusion of the run in order to further establish any decrement in their ability to perform an arduous task at the conclusion of a typical run.

337. The eleven officers that participated in the simulated runs to establish the pace of the officer backup and assist calls were all supervisors. Demonstrating their opposition to improving the fitness of incumbent SEPTA officers, the union did not allow its officers to participate in the simulations. Dr. Moffatt was not concerned that he received supervisory officers to develop the pace because all of the supervisors, like the transit police officers, are held to the same fitness standards.

338. Captain Harold provided the running times and dummy drag times to Dr. Moffatt for use in laboratory testing in Florida.

339. From the simulations in Philadelphia, Dr. Moffatt was able to establish an average assist response pace of 187 seconds. Laboratory simulations were then setup with a tread-

mill and a bench stepping device where Dr. Moffatt could control the work performed and measure the amount of oxygen consumed, as well as the energy expenditure for that work. Dr. Moffatt made sure that the laboratory simulation modeled the concourse that was run in Philadelphia with respect to the distances, angles and number of steps.

340. Dr. Moffatt obtained test participants in Florida to perform the laboratory tests. All of the participants in Florida were tested for their aerobic capacity and for their ability to do anaerobic work such as a dummy drag, a sled pull and an arm crank test. Approximately 95 test subjects participated in the Florida experiments. The Florida test subjects mimicked the age of the average SEPTA officer applicant. The test subjects in Florida ranged in aerobic capacity from approximately 36 mL/kg/min to 58 mL/kg/min.

341. In Dr. Moffatt's first experiment, the participants ran on a treadmill and bench stepping machine and then had their oxygen consumption measured. The run lasted 187 seconds. At the conclusion of the running simulation, each participant performed an arm crank test which approximates a struggle at the conclusion of a run.

342. From the data gathered from the first simulation test, Dr. Moffatt was able to conclude that participants with aerobic capacities of 45 mL/kg/min or better had a nominal decrement in their ability to perform the arm crank simulation to the extent of 9-10%; in contrast, those participants with aerobic capacities of less than 45 mL/kg/min suffered very serious drop-offs in their ability to do work to the extent of a 30% work decrement.

343. Test two was performed in an outdoor setting in Florida. The same distances were run as in the SEPTA concourse and the same numbers of steps were included. Oxygen consumption was then measured. Again, the participants were instructed to perform anaerobic work at the conclusion of the outdoor running test. The conclusions were the same. Those with the aerobic capacities of 45 mL/kg/min or better suffered approximately a 10% decrement in their ability to perform the arm crank test. Those with aerobic capacities of less than 45 mL/kg/min suffered an approximate 30% decrement in their

ability to perform anaerobic tasks at the conclusion of the run. Dr. Moffatt's second experiment reflects that SEPTA transit police officers, after performing a .35 mile run for backup or assist, would suffer similar decrements if they encountered anaerobic tasks, such as an altercation, at the conclusion of the run.

344. Experiment 3 featured the same running protocol but individuals participating in this Florida simulation dragged a dummy for thirty feet at the conclusion of the running portion of the test. Again, individuals with less than 45 mL/kg/min of aerobic capacity suffered decrements approximating 30% while those with aerobic capacities of greater than 45 mL/kg/min suffered decrements of 10-11%.

345. Experiment 4 featured the same running protocol but concluded with a 166 pound sled push that mimics an altercation that could occur at the end of a run at SEPTA. Again, those with aerobic capacities of 45 mL/kg/min or greater suffered an approximate 7-8% decrement in their ability to do anaerobic work while those who scored below 45 mL/kg/min in aerobic capacity had decrements of roughly 30%.

346. Dr. Moffatt also found from his experiment that individuals with an aerobic capacity of less than 45 mL/kg/min had to perform the .35 mile run at between 90% and 95% of their maximum capabilities and sometimes even higher. Those individuals who had aerobic capacities of 45 mL/kg/min or greater were running the .35 mile run at 80% to 85% of their maximum aerobic capacity. In order to determine whether the rate at which a person was running was affecting the results, Dr. Moffatt required individuals with an aerobic capacity of 45 mL/kg/min to run at 95% of their maximal aerobic capacity. Notably, even when the 45 mL/kg/min or higher group was made to run at 95% of their maximal aerobic capacity, their resulting decrement and ability to do anaerobic tasks at the conclusion of the run remained unchanged.

347. From his studies, Dr. Moffatt was able to determine that individuals in the higher fitness group (45 mL/kg/min or higher) have a greater reserve to draw upon at the end of a .35 mile run. Even when operating at close to maximal aerobic

capacity, the higher fitness group has the same ability to draw on their reserve to perform the same amount of anaerobic work at the conclusion of the run.

348. Dr. Moffatt's laboratory experiments were statistically significant at the .05 level.

349. Based on his studies, Dr. Moffatt believes that SEPTA's aerobic capacity standard of 42.5 mL/kg/min as it relates to transit police officer work is very conservative. Indeed, Dr. Moffatt believes that the aerobic capacity cutoff for SEPTA transit police officers should be 45 mL/kg/min.

350. The practical significance of Dr. Moffatt's studies is that a SEPTA transit police officer with an aerobic capacity less than 45 mL/kg/min has to run 3-5 blocks working at maximal effort and may not arrive in a reasonable time period, and if they do arrive in a timely fashion, their ability to do anaerobic work drops off so significantly that they may be ineffective upon arrival.

351. Dr. Moffatt concludes that it would be irresponsible for SEPTA to accept the normative data from the Cooper Institute at the fiftieth percentile for female applicants. As stated above, the fiftieth percentile for women translates into an aerobic capacity of 36 mL/kg/min. Based on Dr. Moffatt's studies, female officers with an aerobic capacity of 36 mL/kg/min would not be able to perform their duties with respect to the amount of work necessary upon arrival after being called in for an assist or backup.

352. Dr. Moffatt also conducted a further experiment comparing groups of high aerobic capacity and high anaerobic capacity persons to a group of low aerobic capacity and high anaerobic capacity persons.

353. This study determined the effect that anaerobic abilities have on work that was being performed at the end of a running test such as a SEPTA backup or assist.

354. Dr. Moffatt concluded that individuals with high aerobic capacities suffered a lesser work decrement than those individuals with lower aerobic capacities, despite the fact that individuals with low aerobic capacities had a very high anaerobic capacity.

355. The group with the high aerobic capacity and high anaerobic capacity had decrements of approximately 10-11%. The group that had high anaerobic capacity but a low aerobic capacity suffered decrements in the vicinity of 25% to 28%.

356. In response to Dr. McArdle's comment that muscular strength and endurance training should specifically train muscles used for a specific position, Dr. Moffatt indicated that individuals require overall strength before they can effectuate a specific technique such as defensive tactics. Moreover, Dr. McArdle's proposed "alternative" does not feature specificity training.

L. The Opinion and Report of Dr. Norman Henderson in Support of the Job-Relatedness and Business Necessity of SEPTA's Physical Fitness Test

357. In support of the job-relatedness and business necessity of its physical fitness test, SEPTA offered the testimony of Dr. Norman Henderson, an industrial and organizational psychologist.

358. Dr. Henderson's report states that "more is better" with regard to muscular strength and endurance vis-a-vis the job duties of a SEPTA transit police officer. Accordingly, Dr. Henderson concludes that SEPTA's gym-based components are job-related.

359. Dr. Henderson has reviewed Dr. Davis' validation study and believes that Dr. Davis has a compelling construct validation argument in his study.

360. To begin, Dr. Henderson states that it was evident that the SEPTA police officer job has a heavy aerobic component in that so many of its critical tasks involve two minutes or more of running. In addition, Dr. Davis also had an enormous body of literature spanning 60 years that demonstrates the linear relationship between aerobic capacity and the ability to do work.

361. Dr. Henderson also noted that Dr. Davis had empirical research supporting his validation study at SEPTA. First, he had a body of empirical research demonstrating that the 1.5 mile run is a valid indicator of aerobic capacity. Dr. Davis also had performed work in other jurisdictions, more particularly

Anne Arundel County, wherein he used a similar job analysis technique and demonstrated empirically that a relationship existed between aerobic capacity and performance tasks in a safety force situation.

362. Dr. Henderson also believes that Dr. Davis' use of the construct validation strategy is in accordance with the SIOP Principles.

363. In sum, Dr. Henderson contends that it was proper for Dr. Davis to create an aerobic capacity test for SEPTA.

364. Dr. Henderson further asserts that it was proper for Dr. Davis to include the gym-based components to measure absolute and relative strength. Indeed, Dr. Henderson was able to validate Dr. Davis' constructs by aggregating the mathematical data and redemonstrating that the absolute and relative standards set by Dr. Davis did correlate with successful job performance by SEPTA police officers.

365. With respect to Drs. Griffin's and Siskin's criterion-related studies, Dr. Henderson submits that the criterion measures used by Drs. Siskin and Griffin — Part I arrests, arrest rates and commendations — are appropriate criterion measures for a transit police force.

366. Dr. Henderson also contends that there has historically been difficulty in using performance evaluations as a criterion for measuring police work due to potential bias that may exist in such subjective evaluations.

367. Dr. Henderson testified that the fact that incumbent transit police officers have failed incumbent aerobic capacity tests or muscular strength and endurance tests is irrelevant to the validity of the test developed as a selection device. Dr. Henderson testified that using incumbents as a benchmark to determine whether a selection device is valid is dangerous for several reasons. Initially, a selection device is not designed to be an absolutely perfect predictor for all members of a company. Also, the incumbent argument incorrectly assumes that the incumbent population will necessarily match the applicant population. Incumbents are generally older individuals than those who a selection device is being used on for new hiring. Moreover, a second fallacious assumption is that the incumbent population is performing well; admittedly, there will be

considerable variation in effectiveness of workers already on a job. Generally, applicants train for a test where incumbents will basically walk in and take a test without any preparation. Therefore, in Dr. Henderson's opinion it is risky to use incumbent data as a benchmark for establishing entry-level selection devices.

M. Alternative Selection Devices

368. During the course of the trial, plaintiffs suggested several alternative selection devices that are allegedly less discriminatory than SEPTA's existing physical fitness test and that would equally serve SEPTA's business interest in having a police officer workforce capable of performing the physical requirements of the job. The tests that plaintiffs propose can be placed into two different groups: (1) no physical fitness testing pre-hire with training to follow and (2) gender-adjusted pre-hire tests with training to follow.

369. Although plaintiffs introduced evidence showing that many law enforcement organizations have no physical entrance requirements pre-hire, the Court will focus in on the Philadelphia Police Department's selection device because this selection device was one of the primary focus points of plaintiffs' alternative selective device argument.

370. The Philadelphia Police Department has no physical entrance requirements. Under this alternative, SEPTA would continue to send its recruits for training to the Academy, where they would be required to pass the Act 120 requirements as established by the Pennsylvania Municipal Police Officers Education and Training Commission standards at the conclusion of their training at the Academy. Officers hired under this standard would only be required to pass the Academy's physical fitness test, which is gender- and age-adjusted.[10]

---

10. A test can be said to be gender- and age-adjusted where a particular component of a test contains different scores for men and women and different scores for different ages. For example, a push-up test that was gender-adjusted may require men between the ages of 20-29 to complete 20 push-ups and women between the ages of 20-29 to complete 15 push-ups. A push-up test that was age-adjusted may require men between the ages of

371. The second alternative to SEPTA's physical fitness test proposed by plaintiffs is to administer a measure of physical fitness as part of the entry-level selection process and provide training to recruits on the specific physical tasks required of SEPTA police officers. Plaintiffs argue that such a selection device would ensure that SEPTA selects officers who have achieved an appropriate level of physical fitness and readiness to complete successfully the physical rigors of the training academy and the physical demands of public safety personnel.

372. Dr. William McArdle has proposed such a test. Dr. McArdle's proposed test evaluates an applicant's general physiologic performance capabilities and readiness to become involved in strenuous physical activity and specific exercise and physical task training that takes place at the Academy and later on the job.

373. Dr. McArdle's proposed test measures the following parameters of physical fitness: (1) lower back and hamstring flexibility (sit-and-reach test); (2) cardiovascular-aerobic fitness (1.5-mile run); (3) abdominal muscular endurance (sit-ups in one minute); (4) upper body muscular strength (1-repetition maximum bench press strength per pound of body weight ratio).

374. Because Dr. McArdle's proposed test measures the applicant's general level of physical fitness, plaintiffs claim that it is proper to recognize the well-established physiological differences between men and women in evaluating the applicant's status for physical fitness. Therefore, the applicant's fitness level is determined by measuring the applicant's score on each of the test's components against standards for those sharing similar immutable sex-specific traits; in essence, the components of McArdle's proposed test would be gender- and age-adjusted, containing different passing scores based on your age and gender.

375. Dr. McArdle's proposed test requires that an applicant achieve a fitness level at the fiftieth percentile for his/her sex on each of the fitness measures based on the normative

---

20-29 to complete 20 push-ups and men between the ages of 30-39 to complete 18 push-ups.

data gathered by the Cooper Institute. For example, on the flexibility component, women must achieve a higher absolute score on the sit-and-reach test than men; for the fiftieth percentile, this equates to a score of 20 inches for women, while men must achieve a sit-and-reach score of only 17.5 inches. This is because empirical data consistently demonstrate that females, as a group, have greater lower back and hamstring flexibility than males. Similarly, each candidate must achieve a physical fitness level at the fiftieth percentile for their sex on the aerobic capacity test. For male candidates, the fiftieth percentile corresponds to a running time of 12:18; for female candidates, the corresponding run time of 14:55 is required.

376. With respect to the gym-based components of SEPTA's physical fitness test, plaintiffs propose that a less discriminatory alternative is the criterion tests (other than the 1.5 mile run) that SEPTA actually adopted in late 1995 or early 1996. According to SEPTA, the criterion tests measure the same thing as the gym-based components. Accordingly, the criterion tests serve the same interest as SEPTA's gym-based components.

377. Plaintiffs also propose that another alternative to the gym-based components of SEPTA's physical fitness test is to have no physical entrance requirements which measure muscular strength and endurance but to provide task-specific training to SEPTA recruits after they are hired. This is the approach used by SEPTA with respect to other requirements of the job, such as the use of firearms, self-defense tactics and effectuating arrests. That is, SEPTA does not require applicants to have firearms training or certification at the time of their application. Rather, applicants acquire this knowledge and ability through training at the Academy. Plaintiffs propose that SEPTA use this same approach for muscular strength and endurance.

378. As will be discussed in greater detail in the Conclusions of Law, not one of plaintiffs' proposed alternative tests is an acceptable alternative selection device under Title VII.

CONCLUSIONS OF LAW

1. Title VII "proscribes not only overt discrimination but also practices that are fair in form, but discriminatory in operation." *Griggs v. Duke Power Co.*, 401 U.S. 424, 431, 91 S. Ct. 849, 853, 28 L. Ed. 2d 158 (1971). Under a disparate impact theory, a showing of discriminatory purpose or intent is not required. *International Bhd. of Teamsters v. United States*, 431 U.S. 324, 335 n.15, 97 S. Ct. 1843, 1854 n.15, 52 L. Ed. 2d 396 (1977).

2. The United States' and the *Lanning* plaintiffs' challenge to SEPTA's physical fitness test for transit police officer applicants is brought under a disparate impact theory. The burdens of proof which are applicable to alleged acts of discrimination occurring on or after November 21, 1991, the effective date of the Civil Rights Act of 1991, are set forth in Section 105 of the Civil Rights Act of 1991, 42 U.S.C. ss. 2000e- 2(k)(1), and applicable case law. These are the burdens of proof applicable to the *Lanning* plaintiffs. They are also the burdens applicable to the United States with respect to all claims of discrimination that occurred after November 21, 1991.

3. With respect to alleged acts of discrimination occurring after November 21, 1991, the plaintiffs have the burden of demonstrating that SEPTA's physical fitness standards have an adverse impact against women. After the plaintiffs make this demonstration, the burden shifts to SEPTA "to demonstrate that the challenged practice is job related for the position in question and consistent with business necessity . . . ." 42 U.S.C. ss. 2000e-2(k)(1)(A)(i).

4. A test is job related "if it measures traits that are significantly related to the applicant's ability to perform the job." *Vulcan Pioneers, Inc. v. New Jersey Dep't of Civil Service*, 625 F. Supp. 527, 545-46 (D.N.J. 1985) (citations omitted), *aff'd*, 832 F.2d 811 (3d Cir. 1987).

5. With respect to SEPTA's second burden of proving business necessity, plaintiffs suggest that the following language from the Supreme Court's opinion in *Dothard v. Rawlinson*, 433 U.S. 321, 331 n.14, 97 S. Ct. 2720, 2728 n.14, 53 L. Ed. 2d 786 (1977) controls:

> "[T]he touchstone is business necessity," *Griggs*, 401 U.S. at 431; a discriminatory employment practice must be shown to be necessary to safe and efficient job performance to survive a Title VII challenge.

Based on this snippet from *Dothard*, plaintiffs submit that SEPTA must demonstrate that its physical fitness test is necessary for safe and efficient job performance to survive plaintiffs' Title VII challenge. The Court, however, finds that plaintiffs have misinterpreted the Supreme Court's standard for business necessity by incorrectly relying on this *dictum* from *Dothard*.

6. *Dothard* invalidated a height and weight requirement for prison guards that disproportionately excluded women applicants and was not proven to be "job-related." 433 U.S. at 332, 97 S. Ct. at 2728. In reaching this conclusion, the Supreme Court required employer-proof identical to that required in its earlier cases: "the employer must meet 'the burden of showing that any given requirement [has] . . . a manifest relationship to the employment in question.' " *Id.* at 329, 97 S. Ct. at 2727 (quoting *Griggs*, 401 U.S. at 432, 91 S. Ct. at 854). Although *Dothard* follows prior Court cases, the Court added the above-quoted footnote language upon which plaintiffs rely for their formulation of the business necessity standard. This footnote formulation, however, is contradicted by the broader standard applied in the *Dothard* text. *Contreras v. City of Los Angeles*, 656 F.2d 1267, 1279 (9th Cir. 1981) (citing *Dothard*, 433 U.S. at 331-32, 97 S. Ct. at 2727-28).

7. After *Dothard*, the Supreme Court has explained that the *Griggs* and *Albemarle Paper Co. v. Moody* standard, rather than the *Dothard* footnote, controls Title VII cases. In *New York City Transit Authority v. Beazer*, 440 U.S. 568, 99 S. Ct. 1355, 59 L. Ed. 2d 587 (1979), the plaintiffs challenged a Transit Authority ("TA") refusal to hire narcotics users, specifically methadone users. The Court stated:

> Respondents recognize, and the findings of the District Court establish, that TA's legitimate employment goals of safety and efficiency require that exclusion of all users of illegal narcotics, barbiturates, and amphetamines, and of

> a majority of all methadone users. The District Court also held that those goals require the exclusion of all methadone users from the 25% of its positions that are "safety sensitive." Finally, the District Court noted that those goals are significantly served by—even if they do not require—TA's rule as it applies to all methadone users including those who are seeking employment in nonsafety- sensitive positions. The record thus demonstrates that TA's rule bears a "manifest relationship to the employment in question." *Griggs*, 401 U.S. at 432, 91 S. Ct. at 854. *See Albemarle Paper Co. v. Moody*, 422 U.S. 405, 95 S. Ct. 2362, 45 L. Ed. 2d 280 (1975).

Id. at 587 n.31, 99 S. Ct. at 1366 n.31.

8. In light of *Beazer*, the Supreme Court's application of the employer's Title VII burden of proof after *Dothard* not only follows the standards set forth in *Griggs* and *Albemarle*, but implicitly approves employment practices that significantly serve, but are neither required by nor necessary to, the employer's legitimate business interests. Thus, to demonstrate business necessity, SEPTA need only show that the 1.5 mile run component of its physical fitness test bears a manifest relationship to the position of SEPTA transit police officer.

9. Specifically, as the Ninth Circuit found in *Contreras*, this Court finds that "discriminatory tests are impermissible unless shown, by professionally accepted methods, to be predictive or significantly correlated with important elements of work behavior that comprise or are relevant to the job or jobs for which candidates are being evaluated." 656 F.2d at 1280.

10. If SEPTA satisfies its burden of persuasion, the United States may still prevail if it demonstrates that an alternative employment practice has less disparate impact and "would also serve the employer's legitimate interest in 'efficient and trustworthy workmanship.'" *Albemarle*, 422 U.S. at 425, 95 S. Ct. at 2375. That is, the United States may prevail if it demonstrates that the alternative test would "be equally as effective as the challenged practice in serving the employer's legitimate business goals." *Watson v. Fort Worth Bank & Trust Co.*, 487 U.S. 977, 998, 108 S. Ct. 2777, 2790, 101 L. Ed. 2d 827 (1988).

11. The burdens of proof for alleged acts of discrimination occurring prior to November 21, 1991, are those set forth in *Griggs*, 401 U.S. at 432, 91 S. Ct. at 854, and *Wards Cove Packing Co. v. Atonio*, 490 U.S. 642, 661, 109 S. Ct. 2115, 2127, 104 L. Ed. 2d 733 (1989).[11] Under Wards Cove, after the plaintiffs have made a showing of disparate impact, the burden of production, rather than the burden of persuasion, shifts to the employer to offer evidence of job-relatedness and business necessity. However, the burden of persuasion as to these issues remains with the plaintiffs. The third prong of the disparate impact analysis is still controlled by Albemarle and Watson. Because the Lanning class does not include persons who were rejected by SEPTA prior to 1993, only the United States is seeking relief for persons rejected by SEPTA prior to November 21, 1991.

A. Adverse Impact of 1.5 Mile Run

12. In disparate impact cases such as this, statistical evidence is typically used to establish the adverse impact of an employee selection device under Title VII. SEPTA has admitted that the disparity between the pass rate for male and female applicants on the 1.5 mile run at all times exceeded 2 or 3 standard deviations as measured by the formula set forth in *Hazelwood*, 433 U.S. at 308, n.14, 97 S. Ct. at 2742 n.14, and *Castaneda*, 430 U.S. at 496-97 & n.17, 97 S. Ct. at 1281-82 & n.17 (hereinafter "*Hazelwood* formula"). Indeed, the disparities between the pass rates for male and female applicants for the total of the years 1991, 1993 and 1996 are very large — 5.56 standard deviations — indicating severe adverse impact.

13. These disparities constitute "gross disparities" sufficient to make out a *prima facie* case of discrimination. *Hazelwood*, 433 U.S. at 308 n.14, 97 S. Ct. at 2742 n.14 (disparities larger than two or three standard deviations are generally sufficient to establish a *prima facie* case of discrimination) (citing *Castaneda*, 430 U.S. at 497 n.17, 97 S. Ct. at 1281 n.17); *Teamsters*, 431 U.S. at 340 n.20, 97 S. Ct. at 1857 n.20.

---

11. The burdens of proof as articulated in *Wards Cove* were legislatively overruled by Section 105 of the Civil Rights Act of 1991.

14. In other words, the disparity did not occur by chance and there is a specific cause for the disparity, *i.e.*, discrimination. *See EEOC v. American National Bank*, 652 F.2d 1176, 1192 (4th Cir. 1981) (the *Hazelwood* analysis is utilized by the courts "absolutely to exclude chance as a hypothesis, hence absolutely to confirm the legitimacy of an inference of discrimination").

15. Moreover, the p-values relating to the 1991, 1993, and 1996 administrations of the 1.5 mile run are .0001, .0001, and .00001 respectively and the p-value relating to the aggregate of these three administrations of the 1.5 mile run is .00001.

16. Numerous courts have accepted that a p-value below .05 indicates that a difference is statistically significant. *See, e.g., Bouman v. Block*, 940 F.2d 1211, 1225-26 & n.1 (9th Cir. 1991); *Palmer v. Shultz*, 815 F.2d 84, 92 & 96 (D.C. Cir. 1987).

17. Accordingly, the Court concludes that the 1.5 mile run requirement of SEPTA's physical fitness test has a severe adverse impact against women.[12]

B. The Job-Relatedness and Business Necessity of The 1.5 Mile Run Component of SEPTA's Physical Fitness Test

18. With respect to administrations of SEPTA's physical fitness test after November 21, 1991, the effective date of the Civil Rights Act of 1991, the Court concludes that SEPTA has established the job-relatedness and business necessity of the 12 minute, 1.5 mile run component.

19. With respect to the administrations of SEPTA's physical fitness test prior to November 21, 1991, the effective date of the Civil Rights Act of 1991, the Court concludes that the United States has failed to demonstrate that the 12 minute, 1.5 mile run component of SEPTA's physical fitness is not job-related or not consistent with business necessity.

20. In sum, the Court concludes that SEPTA's proffered evidence of validity does establish the job-relatedness and business necessity of the 12 minute, 1.5 mile run requirement of SEPTA's physical fitness test.

---

12. The Court will address below the United States' challenge to the gym-based components of SEPTA's physical fitness test.

21. Studies done post hoc in an attempt to validate a test already given and in anticipation of litigation must be carefully scrutinized due to a danger of lack of objectivity. *See Albemarle*, 422 U.S. at 433 n.32, 95 S. Ct. at 2379 n.32.

22. As stated above, the employer's burden of establishing the validity of a selection device requires a showing that the challenged device has a "manifest relationship to the employment in question." *Griggs*, 401 U.S. at 432, 91 S. Ct. at 854. Proof that an examination is job-related must be based on a study that meets "professionally acceptable" standards and procedures. *Albemarle*, 422 U.S. at 431, 95 S. Ct. at 2378. Specifically, the Supreme Court has stated that examinations that have a significant adverse impact upon protected groups are impermissible unless shown, by professionally acceptable methods, to be "predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated." *Id.* at 431, 95 S. Ct. at 2377 (citation omitted).

23. Although the Supreme Court has instructed that the *Uniform Guidelines* are "entitled to great deference," *id.* at 431, 95 S. Ct. at 2378; *Griggs*, 401 U.S. at 433, 91 S. Ct. at 855, the Court has retreated subsequently from its strict adherence to the *Uniform Guidelines*. "In *Washington v. Davis*, [426 U.S. 229, 96 S. Ct. 2040, 48 L. Ed. 2d 597 (1976),] the Court concluded that a test shown to successfully predict performance in police training was justified despite the fact that neither the test nor the training program had been validated as predictors of job performance, as required by the guidelines." Paul N. Cox, Employment Discrimination P. 12.03 (2d ed. 1992) (citing *Washington*, 426 U.S. 229).[13]

In *Watson*, the Supreme Court explicitly held that "[o]ur cases make it clear that employers are not required, even when defending standardized or objective tests, to introduce formal 'validation studies' showing that particular criteria predict actual on-the-job performance." *Watson*, 487 U.S. at 997, 108 S. Ct. at 2791 (citing *Beazer*, 440 U.S. 568, 99 S. Ct. 1355;

13. Although *Davis* did not involve Title VII directly, the Court applied and interpreted Title VII standards.

Washington, 426 U.S. 229). In light of these more recent Supreme Court cases, it is obvious that an employer's selection device will not necessarily be found to be not job-related or lacking in business necessity due to the fact that the selection device does not strictly adhere to the *Uniform Guidelines*. This result plainly flows from the Supreme Court's holding that an employer need not introduce formal validation studies in support of their employment selection tests; if an employer need not even introduce a validation study, then surely the employer need not comply with every technical requirement of the *Uniform Guidelines* if the employer decides to introduce formal validation studies.

24. The Supreme Court's recent inclination not to require formal validation studies in support of employment tests is logical in light of the fact that many industrial psychologists believe that the *Uniform Guidelines* "cannot be satisfied in practice and incorporate notions originally viewed by the psychologists only as abstract objectives, not as hard and fast criteria of test validation." Cox, *supra*, at P. 12.03 (footnote and citations omitted). Indeed, both plaintiffs' and defendant's experts, here, acknowledge that most employment tests can never be fully reconciled with the *Uniform Guidelines*, and yet these same tests are considered to be professionally acceptable.

25. In light of the foregoing observations, the Court finds that SEPTA's validation tests do not have to satisfy every intricate detail of the *Uniform Guidelines* to be considered professionally acceptable. Instead, SEPTA merely has to demonstrate, by professionally acceptable methods, that the 1.5 mile run of its physical fitness test is "predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated." *Albemarle*, 422 U.S. at 431, 95 S. Ct. at 2377.

26. In support of the job-relatedness and business necessity of the 1.5 mile run component of the physical fitness test, SEPTA has offered Dr. Davis' validation study, the criterion-related validation studies of Drs. Griffin and Siskin, the study of Dr. Moffatt and the testimony and report of Dr. Henderson.

As will be discussed in greater detail below, the Court finds that this evidence establishes the job-relatedness and business necessity of the 1.5 mile run.

27. The Court first rejects plaintiffs' suggestion that aerobic capacity is not a required physical ability for a SEPTA transit police officer. Based on all of the evidence presented at trial, the Court concludes that the predominant energy system utilized by SEPTA transit police officers during the course of their duties, especially during pursuits, officer backups and officer assists, is aerobic metabolism. The Court credits the testimony of Drs. Moffatt and Davis that aerobic capacity is the primary and predominant source of energy supporting the SEPTA foot-based transit patrol force. The Court specifically finds Dr. McArdle's testimony - that anaerobic energy is the predominant energy source for SEPTA transit police officers - not credible; his testimony contradicts the patrol officer testimony and is inconsistent with the described lengths and durations of jogging, sprinting, and running activities carried out by SEPTA patrol officers on a daily basis.

28. The Court finds that Dr. Davis' validation study, which utilized a construct validity strategy, has sufficient empirical support for the aerobic capacity requirement of 42.5 mL/kg/min. Given the frequency of jogging, sprinting, running, stair climbing of considerable heights and other such arduous tasks required of SEPTA officers, the aerobic capacity of 42.5 mL/kg/min was readily justifiable. However, the Court notes that Dr. Davis did not rely on judgment alone. Rather, Dr. Davis had empirically established in Anne Arundel County that an aerobic capacity equal to 42.5 mL/kg/min predicted successful performance on a police officer work sample test.

29. Dr. Davis' decision to require 42.5 mL/kg/min of aerobic capacity was supported both empirically and by his considerable experience in developing tests for law enforcement agencies. For these reasons, the Court finds that Dr. Davis' study, standing alone, met the professional standards for construct validation and satisfies defendant's burden of demonstrating job relatedness and business necessity. As the SIOP Principles acknowledge:

> *[j]udgment is necessary in setting any critical or cutoff score. A fully defensible empirical basis for setting a critical score is seldom, if ever, available.* The only justification that can be demanded is that critical scores be determined on the basis of a rationale which may include such factors as estimated cost-benefit ratio, number of openings and selection ratio, success ratio, social policies of the organization, or judgments as to require knowledge, skill or ability on the job. If critical scores are used as a basis for rejecting applicants, their rational or justification should be made known to the users.

SIOP Principles at 32-22 (emphasis added). Dr. Davis' validation study satisfies this standard in that it articulates a justification for using a cutoff score of 42.5 mL/kg/min on SEPTA's physical fitness test.

30. Plaintiffs argue that Dr. Davis' study cannot be found to support the job-relatedness or business necessity of the 1.5 mile run component of SEPTA's physical fitness test because the study did not satisfy all of the technical requirements of the *Uniform Guidelines*. While plaintiffs correctly contend that Dr. Davis' study does not satisfy all of the standards of the *Uniform Guidelines*, plaintiffs incorrectly conclude that these violations undermine the overall validity of Dr. Davis' study. As stated above, the Supreme Court does not require employers, even when defending standardized or objective tests, to introduce formal "validation studies" showing that particular criteria predict actual on-the-job performance. Consequently, it is irrelevant that Dr. Davis may not have strictly complied with all of the technical standards of the *Uniform Guidelines*, rather what is important is that Dr. Davis' study meets professionally acceptable standards.

31. In light of all of the evidence introduced at trial, the Court concludes that Dr. Davis' study meets professionally acceptable standards. Dr. Henderson, who is an expert in the development of physical abilities tests, specifically testified that Dr. Davis' study constitutes a proper construct validity study. Further, an independent review of the evidence establishes that Dr. Davis' study constitutes a proper construct validity study. The evidence demonstrates that Dr. Davis'

study identifies essential and critical physical tasks required of SEPTA transit police officers that require a high aerobic capacity. In addition, Dr. Davis' study establishes that the 1.5 mile run tests for the trait of aerobic capacity, *i.e.*, endurance, stamina and cardiovascular reserve, which is necessary to the performance of various physical tasks encountered by SEPTA officers.

32. During the trial, plaintiffs' expert, Dr. Zedeck, criticized Dr. Davis' study for failing to comply with the *Uniform Guidelines* in some instances. However, on cross- examination, Dr. Zedeck revealed that a physical abilities test he created for the San Francisco Fire Department suffered from many of the same deficiencies that he found in Dr. Davis' study. Despite these deficiencies, Dr. Zedeck testified that his study was properly validated. In essence, Dr. Zedeck implicitly conceded, through his admission that his study was properly validated despite its errors, that Dr. Davis' study could be considered properly validated even though that study was not fully defensible.

33. The Court agrees with Dr. Zedeck's implicit admission that Dr. Davis' study can be considered properly validated even though it is not free from errors. Admittedly, validation studies by their very nature are "difficult, expensive, time-consuming and are rarely, if ever, free of errors." *See Cleghorn v. Herrington*, 813 F.2d 992, 996 (9th Cir. 1987). Thus, it is irrelevant that Dr. Davis' study may have errors in light of the fact that psychological experts, case law and the SIOP Principles all recognize that no studies will ever be without errors. Instead, the more appropriate question is whether Dr. Davis' study comports with professionally acceptable standards, and the Court finds that this question can be answered in the affirmative.

34. As remarked earlier, test validation attempts to determine whether (and the degree to which) persons who are selected by a test will be successful performers on the job, and whether those who are not selected would not have been successful performers on the job. Dr. Davis' test achieves this objective. Plainly, it is more likely than not that applicants who pass the 1.5 mile run component of SEPTA's physical fitness

test will be successful performers on the job; whereas, it is highly probable that those officers who do not pass the 1.5 mile run component of SEPTA's test will not be successful performers on the job because they lack the aerobic capacity necessary to fulfill the demanding obligations of a SEPTA officer.

35. In addition to the Court's findings relating to Dr. Davis' study, the Court finds that the continuing validation studies of defendant's experts, Drs. Griffin and Siskin, also demonstrate the job-relatedness and business necessity of the 1.5 mile run component of SEPTA's physical fitness test.

36. Dr. Siskin, with the assistance of Dr. Griffin, conducted several studies on behalf of SEPTA to determine whether SEPTA's requirement of 42.5 mL/kg/min of aerobic capacity predicted patrol officer performance in the areas of arrests, including Part I crimes and overall arrests. Dr. Siskin also was asked to determine whether SEPTA's aerobic capacity requirement predicted or correlated with the receipt of commendations for patrol officer activities that concerned "street performance" involving arrests. Further, Dr. Siskin tabulated the aerobic capacity of the SEPTA transit police force and compared the aerobic capacity of the SEPTA transit police force with what he estimated to be the aerobic capacity of perpetrators of crimes within the SEPTA transit system.

37. Dr. Siskin found that SEPTA officers with an aerobic capacity of 42 mL/kg/min or higher had statistically significant higher rates of actual arrests with respect to Part I offenses, higher rates of overall arrests and higher Part I arrest rates when compared to officers who were below SEPTA's aerobic capacity requirement of 42 mL/kg/min. In sum, officers who met or exceeded SEPTA's aerobic capacity requirement made more arrests, particularly Part I arrests, than those officers who had an aerobic capacity below SEPTA's requirement of 42 mL/kg/min and were more likely to make both Part I arrests and overall arrests than those officers below 42 mL/kg/min.

38. Dr. Siskin also determined that a linear relationship existed between aerobic capacity and arrests and arrest rates.

This linear relationship demonstrates that the higher the aerobic capacity of the officer, the higher the officer's arrest rate, number of Part I arrests and overall arrests. These findings were statistically significant at less than .05 and in many cases less than .001, thus meeting the significance requirements (.05) of the *Uniform Guidelines.*

39. Dr. Siskin also found that the likelihood of receiving a commendation for "street" patrol officer performance was statistically significantly higher if the officers' aerobic capacity met or exceeded 42 mL/kg/min. Dr. Siskin reviewed 207 commendations that were awarded for the period of 1994 through 1996 and found that 96% of the commendations went to officers who had an aerobic capacity greater than 42 mL/kg/min; these officers had an average aerobic capacity of 46 mL/kg/min. Further, 198 of the commendations studied involved an arrest with 116 having an explicit reference in the commendation document to a foot pursuit, use of force or other physical exertion.

40. Dr. Siskin's testimony further established that, when comparing officers who were *always* at 42 mL/kg/min or over to officers who were always under 42 mL/kg/min, the higher aerobic capacity group had a 57.1% "arrest rate" advantage in the more serious Part I crimes and 28% greater arrest rate for all offenses. Dr. Siskin's data also showed that officers always at 42 mL/kg/min or above made three times (151%) the *actual* number of Part I arrests and 75% more *actual* overall arrests when compared to officers who never met the 42 mL/kg/min standard.

41. Plaintiffs challenge the reliability of Dr. Siskin's studies by noting that certain "contaminating factors"[14] may have affected the results of these studies. Notwithstanding plaintiffs' contentions, the Court finds that Dr. Siskin credibly addressed plaintiffs' concerns about the alleged contaminating factors, including age, tenure and learning, by controlling for rank and assignment. Dr. Siskin did this through a "regression

14. A contaminating factor can be described simply as a factor that possibly affects the results of an observed statistical relationship. Plaintiffs raised the issue of contaminating factors in order to create doubt as to whether the statistical relationships observed by Dr. Siskin were accurate.

analysis" that adjusted for zone, shift and rank. The regression analysis allows the Court to compare officers who are similarly situated with respect to their assignments.

42. The regression analysis showed that the differences between the officers who achieved 42 mL/kg/min or higher versus the officers who never met 42 mL/kg/min was still statistically significant in the number of Part I arrests made, the arrest rate for Part I crimes and the arrest rates for all crimes, comparing officers of the same rank who were assigned to the same zone and tour. Specifically, after the regression analysis was run, Dr. Siskin's data showed: a 14% advantage in the overall arrest rate for officers at or above 42 mL/kg/min and a 32% arrest rate advantage for officers at or above 42 mL/kg/min for Part I crimes, as well as a significant difference in the number of Part I arrests made by officers meeting or exceeding SEPTA's aerobic capacity standard.

43. Dr. Siskin's testimony also established that rotating officers through different beats would have no effect on his conclusions because beat assignments are not correlated to an officer's aerobic capacity.

44. Dr. Siskin's testimony also established that beat assignments are simply random "noise" that obscures and lowers the observed correlation coefficients and statistical significance. Notwithstanding this noise, all of Dr. Siskin's studies were statistically significant at either less than the .05 level or less than the .01 level, and in many instances less than the .001 level. Dr. Siskin testified that running a partial correlation for beat assignments would have only raised the correlation and the level of statistical significance.

45. Dr. Siskin explained that random errors in measurement or errors in the data can be considered the same as random noise, that is, Dr. Siskin testified that there was no reason to believe that these types of errors - measurement, data and attribution errors - will favor either a high aerobic capacity group or low aerobic capacity group; hence, they are random with respect to aerobic capacity and act as random noise. In essence, they have no effect on the observed statistical relationships.

46. The testimony of Dr. Siskin showed that random noise would only suppress correlations once a significant relationship between aerobic capacity and the arrest parameters has been observed. Random noise or random errors cannot create a statistical relationship; indeed, such randomness only masks such a relationship. Dr. Siskin further explained that once a correlation is observed and adjustments are made for random noise or error, the statistical corrections will raise the correlation. Consequently, Dr. Siskin found that the observed correlations in this case were an underestimation of the true relationship between satisfying SEPTA's aerobic capacity requirement and making Part I arrests, overall arrests and arrest rates.

47. Dr. Siskin's testimony demonstrated that while corrections for random noise would "clearly increase the correlations" so that the estimates of the correlations that he obtained were actually too low, he did not make these corrections because the best measure of practical significance is found through regression analysis and expectancy tables, which estimate the effect of meeting SEPTA's aerobic capacity standard (42 mL/kg/min) relative to not meeting SEPTA's standard. More significantly, these estimates are unaffected by random noise.

48. The import of Dr. Siskin's testimony was that once a relationship between aerobic capacity and arrest and arrest rates was found in the data, any controls for random noise, measurement errors or any other factors random with respect to an officer's aerobic capacity level would only have raised the correlation and increased the statistical significance which was already at less than .05 and less than .01 levels.

49. In sum, the evidence establishes that there exists a statistically significant correlation between the 1.5 mile run component of SEPTA's physical fitness test and SEPTA's objective measures of job performance, such as arrest rates, arrests and commendations, and importantly, this evidence was never refuted by plaintiffs.

50. Plaintiffs, however, argue that Dr. Siskin's studies cannot support the job-relatedness or business necessity of the 1.5 mile run component of SEPTA's physical fitness test because

the observed correlation coefficients do not exceed + .30 on the officer basis. *See Hamer v. City of Atlanta*, 872 F.2d 1521, 1525-26 (11th Cir. 1989); *Dickerson v. United States Steel Corp.*, 472 F. Supp. 1304 (E.D. Pa. 1978). Plaintiffs' argument, however, proceeds on the faulty assumption that practical significance can only be measured by correlation coefficients and that these correlation coefficients must exceed + .30 to have any legal effect. As will be explained, the case law does not mandate that practical significance must be shown only by correlation coefficients, and more importantly, industrial and organizational psychologists recognize that practical significance may be more properly demonstrated by examining regression analyses and expectancy tables.

51. Dr. Siskin testified that the best indicator of the practical significance of aerobic capacity is determined through regression analysis which explicitly measures the expected gain in arrests resulting from the aerobic capacity standard of 42 mL/kg/min. Dr. Siskin demonstrated through a regression analysis that SEPTA could have achieved 470 additional overall arrests — 70 of which were Part I arrests for serious crimes for the period of 1991 through 1996.[15] These findings reflect a 10% increase in Part I arrests and a 4% increase in the overall arrest rate. The practical significance analysis included a regression that took into account all relevant variables, including rank, zone, tour and unfounded incidents, and also controlled for special units. Dr. Siskin testified that taking these variables into account, the statistical relationship and predictive nature of aerobic capacity remained the same, thus demonstrating that meeting SEPTA's aerobic capacity standard of 42 mL/kg/min consistently predicted higher Part I arrests and higher arrest rates for all crimes and Part I crimes for officers who maintained higher levels of aerobic capacity when compared to those officers that failed to meet SEPTA's aerobic capacity standard of 42 mL/kg/min.

---

15. Dr. Siskin found that SEPTA could expect a half percent increase in Part I arrests for every increase in mL/kg/min of aerobic capacity and that such an effect was linear.

52. The Court agrees with Dr. Siskin that the correlation coefficient[16] issue is in some sense a "red herring" because an examination of the correlation coefficients does not necessarily explain the practical impact of SEPTA's aerobic capacity requirement. By looking at the correlation coefficients, the Court cannot properly determine how the SEPTA Transit Police Department would benefit from having all of its officers have an aerobic capacity of 42 mL/kg/min or better. However, by looking at Dr. Siskin's regression analysis, the practical significance can be more readily ascertained.

53. In support of his contention that the Court should not look at correlation coefficients to determine the practical significance of the observed correlations between the running test and job performance at SEPTA, Dr. Siskin stated that it is well known and can be proven mathematically that if you are measuring the utility of tests, correlation coefficients are an inappropriate measure. Additionally, Dr. Henderson, a well-qualified expert in tests and measurements for police and fire organizations, indicated that the commonly held view of psychometricians is that the correlation coefficient statistic has no meaning in terms of practical significance.

54. The Court has also reviewed the SIOP Principles and notes that this authoritative treatise makes it clear that correlation coefficients are not the only manner by which practical significance can be determined. Indeed, the SIOP Principles specifically provide that the use of the slope of the regression line or expectancy tables may be the preferred methods in order to determine the practical significance of the test at issue. Indeed, the SIOP Principles specifically recommend the method of determining practical significance that SEPTA has utilized in this case:

> The analysis should provide information about the strength of the relationship, usually a coefficient of correlation. Other methods (such as the slope of the regression line, expectancy tables, or the percentage of misclassifications) are acceptable and may be preferable in many situ-

16. Simply stated, a correlation coefficient establishes the degree of a correlation.

ations. *The analysis should also give information about the nature of the relationship and how it might be used in prediction.*

SIOP Principles at 15 (emphasis added).

55. Based on these foregoing observations regarding practical significance, the Court concludes that correlation coefficients are not the appropriate method to determine the practical significance of SEPTA's aerobic capacity standard. The proper method, in this case, is to use Dr. Siskin's regression analysis that will estimate the expected gain in arrests if officers below 42 mL/kg/min maintained an aerobic capacity of 42 mL/kg/min during the time period in question. In this regard, maintaining such a standard would have resulted in a 10% increase in Part I arrests - an additional 70 Part I arrests - and in a 4% increase in overall arrests - approximately 470 additional arrests.

56. This Court is not unmindful of the significance of the additional 470 overall arrests and additional 70 Part I arrests that would be obtained if SEPTA's less-fit officers met SEPTA's aerobic capacity standard. For many of the 470 additional arrests, there would be fewer criminals in the SEPTA transit system left to prey on and victimize the riding public. Significant gains in apprehensions and deterrence such as those demonstrated here are to be encouraged and supported by the federal courts. The Court simply will not condone dilution of readily obtainable physical abilities standards that serve to protect the public safety in order to allow unfit candidates, whether they are male or female, to become SEPTA transit police officers.

57. Assuming that SEPTA must demonstrate practical significance through correlation coefficients as a matter of law, which the Court does not hold, the Court rejects plaintiffs' assertion that SEPTA has not met its burden of meeting "job-relatedness" and "business necessity" because the correlation coefficients presented in support of Dr. Siskin's study are "low." The Court first notes that the *Uniform Guidelines* only require that the correlations be statistically significant. In this case, each correlation reported by SEPTA was statistically significant at the less than .05 level or less than the .01 level of

statistical significance, and thus are well within the level of significance required by the *Uniform Guidelines*. Indeed, plaintiffs' expert, Dr. Zedeck, admitted that defendant's correlations are statistically significant, especially with respect to the predictive relationship of aerobic capacity and making Part I arrests.

58. Defendant's correlation coefficients are adequate to demonstrate practical significance. Dr. Zedeck testified that he would look to the officer basis data to determine whether plaintiffs have demonstrated practical significance. Applying the officer basis, Dr. Siskin observed a correlation of + .22, which was uncorrected for restriction in range. However, if the officer basis correlation coefficient was corrected for restriction of range, it would reach the magnitude of + .33. *See Bernard v. Gulf Oil Corp.*, 890 F.2d 735 (5th Cir. 1989) (.22 uncorrected correlation coefficient sufficient to demonstrate job relatedness and business necessity); *Commonwealth of Pennsylvania v. O'Neill*, 465 F. Supp. 451, 461, 464-65 (E.D. Pa. 1979) (corrected correlation coefficient of .268 sufficient in police officer case). This + .33 correlation coefficient satisfies the + .30 standard that plaintiffs suggest has to be satisfied in order to show practical significance. Thus, SEPTA has satisfied plaintiffs' absolute standard of practical significance.

59. Using the test event basis of Dr. Siskin's studies, the correlations between meeting SEPTA's aerobic capacity standard and increased levels of Part I arrests and higher arrest rates for serious crimes was + .12. Although this correlation is below + .30, the Court still finds it sufficient to establish practical significance in light of the *Uniform Guidelines* understanding that "there are no minimum correlation coefficients applicable to all employment situations" and in light of Dr. Siskin's testimony that a low correlation coefficient must be expected due to the fact that his studies dealt with low numbers, such as number of officers and number of test events.[17] *See Bernard*, 890 F.2d 735 (recognizing that Supreme Court

17. Dr. Siskin also opined that this correlation was quite strong in light of the fact that he was dealing with such low numbers in terms of number officers and test events.

precedent does not require a minimum cutoff point for correlation coefficients, the court declines to establish bright line cutoff point for correlation coefficient).

60. Utilizing either a correlation coefficient analysis or a regression analysis and expectancy tables, the Court finds that Dr. Siskin's studies have more than amply demonstrated practical significance. Consequently, SEPTA has met its burden in demonstrating that its aerobic capacity test is "predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job" of SEPTA transit police officer. *Contreras*, 656 F.2d at 1283. Thus, the Court finds that SEPTA, through Drs. Siskin's and Griffin's studies, has established the job-relatedness and business necessity of its aerobic capacity requirement.

61. The Court also concludes that SEPTA's "perpetrator analysis" supports its aerobic capacity requirement. This analysis demonstrates that the perpetrator population is approximately 26 years of age and maintains an average aerobic capacity of 47-48 mL/kg/min.

62. It is obvious to this Court that SEPTA transit police officers are frequently required to pursue young male perpetrators that, on average, maintain a high level of aerobic capacity. Dr. Siskin's study of the perpetrator population established that the mean aerobic capacity of the officers receiving commendations was 46.6 mL/kg/min and that the mean aerobic capacity of the officers arresting the perpetrators was 46.8 mL/kg/min. In essence, the perpetrators with high levels of aerobic capacity were being arrested by SEPTA officers with high levels of aerobic capacity, thus lending weight to SEPTA's argument that aerobic capacity is required of SEPTA officers in order to perform successfully on the job.

63. The link between higher levels of aerobic capacity and apprehension of perpetrators is clear to the Court. Further, the plaintiffs' complaints about the perpetrator study, *i.e.*, unsupported inquiries as to purported potential drug and alcohol use by criminals, is dismissed by this Court. This spurious challenge was mere speculation and fails to meet the legal requirements to demonstrate statistically that the inferences drawn from defendant's perpetrator studies are incorrect.

64. The Court also rejects plaintiffs' argument that the perpetrator analysis is irrelevant because it assumes that police work is an "athletic contest." In essence, plaintiffs argue that the perpetrator analysis is flawed because it does not demonstrate that officers with high aerobic capacity are needed to arrest perpetrators with high levels of aerobic capacity. The Court rejects this criticism as contrary to plain common sense. Although the perpetrator analysis shows that SEPTA officers on average have a lower aerobic capacity than the perpetrators arrested for Part I crimes, this same analysis shows that the SEPTA officers who made these arrests have a high level of aerobic capacity, thus establishing that SEPTA officers with high aerobic capacity make more arrests than officers with low aerobic capacity.

65. Moreover, to the extent that potential perpetrators in the SEPTA system have high levels of aerobic capacity, the Court finds that it would be helpful to the successful performance of a SEPTA police officer if the officer also had a high level of aerobic capacity. Although SEPTA officers can use other methods to make an arrest, such as negotiation, officer backup, display of weapon, etc., many situations will arise whereby a SEPTA officer will have to use his aerobic capacity to successfully effectuate an arrest or perform another aspect of his job against a perpetrator who has a high aerobic capacity; therefore, it is beyond cavil that SEPTA officers, if possible, should be as physically fit as, if not more fit, than the perpetrators.[18] Consequently, the Court finds that the perpe-

---

18. During trial, plaintiffs suggested that SEPTA's physical fitness test may not be job-related or consistent with business necessity due to the fact that physical fitness was only one trait required of SEPTA officers. The Court, however, rejects this argument on its face. While it may be true that physical fitness is only one trait or ability required of SEPTA officers, it is a trait or ability that it necessary for and critical to the successful performance of the job, and thus SEPTA should be able to test for such a trait. To suggest otherwise, one would have to ignore common sense and reality. Taking plaintiffs' argument to its logical conclusion, an employer would never be able to test for a particular trait or ability whenever the employment position required many traits or abilities. Of course, Title VII does not impose this prohibition on employers, and plaintiffs are wrong to insinuate that it should. If the position of SEPTA transit police officer requires other abili-

trator analysis supports SEPTA's argument that its aerobic capacity requirement is job-related and consistent with business necessity.

66. To the extent that plaintiffs claim that the statistical analyses offered by SEPTA did not control for various factors and that all the potential variables were not examined, the Court rejects this argument and finds that the relevant and probative variables were controlled for by SEPTA. Furthermore, no statistical study has been offered by plaintiffs that refutes SEPTA's statistical evidence. The plaintiffs have not demonstrated that any of the claimed variables or factors that they assert should have been studied would have made any significant difference in the outcome of SEPTA's studies. Moreover, it is not permissible or appropriate for a party to challenge a regression analysis without proving that the omitted factors would have made a significant difference. *See Bazemore v. Friday*, 478 U.S. 385, 404 106 S.Ct. 3000, 3010 (1986); *Sobel v. Yeshiva University*, 839 F.2d 18 (2nd Cir. 1988); *EEOC v. General Telephone Co. of Northwest, Inc.*, 885 F.2d 575 (9th Cir. 1989).

67. Since *Bazemore*, courts have held that more is required than simply pointing out "potential" flaws in a proponent's statistical analysis in order to rebut the inferences raised by the statistics. A party opposing statistics must do more than simply challenge a proponent's regression studies on the speculative basis that the results might have been different if some unaccounted factor had been included. *See Rossini v. Ogilvy & Mather, Inc.*, 798 F.2d 590, 604 (2nd Cir. 1986). The burden is on the challenger to show how the alleged flaws biased the result. *See General Telephone*, 885 F.2d 575. Here, plaintiffs have not met that burden, rather they have merely speculated as to uncontrolled variables that "may" have affected the results.

---

ties such as negotiation skills (which the Court finds that it probably does), then SEPTA should also be permitted to test for these skills, instead of being precluded from testing for physical fitness as plaintiffs suggest. Thus, to the extent plaintiffs insinuate that SEPTA cannot test for physical abilities because SEPTA officers rely on other skills or abilities to successfully perform their job, the Court rejects this argument.

68. In contrast, the Court finds that SEPTA controlled for rank, tour, zone and unfounded incidents, as well for special units, more than adequately addressing the variables that could have influenced the outcome of any of the studies performed by Dr. Siskin.

69. Plaintiffs also question the appropriateness of using arrests, arrests rates and commendations as criterion measures. The Court, however, finds that SEPTA's use of these arrest criteria and commendations is reasonable and that these criteria are objective criteria upon which Dr. Siskin can properly base his criterion-related validity studies. Indeed, Dr. Henderson testified there is ample support for the use of these criteria from the Law Enforcement Assistant Administration studies and Chicago Police Department studies which have identified certain objective "crime fighting" criteria; the crime fighting criteria include misdemeanor arrests, felony arrests, commendations, court cases and conviction rates. Here, the Court finds that SEPTA's use of three of these objective criteria satisfies the requirements of the *Uniform Guidelines* on the selection of criteria for use in criterion-related validity studies. *See* 29 C.F.R. § 1607.14(B)(3) ("certain criteria may be used without a full job analysis if the user can show the importance of the criteria to the particular employment context"). Additionally, the Court notes that certain witnesses — Chief Evans, Captain Harold, Inspector Pryor of the Philadelphia Police Department and Chief McDevitt from the Washington D.C. Metropolitan Area Transit Authority — all testified that arrests are a critical job task and a valid measure of officer performance.

70. Furthermore, although plaintiffs propose that SEPTA should have used performance evaluations as criterion measures, nothing in the case law, the *Uniform Guidelines* or the SIOP Principles preclude SEPTA from using objective criteria over potentially biased and subjective evaluations. In addition, the Court finds that SEPTA's choice of three of the crime fighting criteria more than adequately supports its validity studies. Moreover, the Court notes that no performance evaluations exist for SEPTA transit patrol officers and that supervi-

sors were evaluated only on their administrative skills. Therefore, performance evaluations in this case are non-existent, and thus irrelevant.

71. The Court also credits Lt. Maslin's testimony that the daily control log - a document that was used as an underlying data source for Dr. Siskin's studies - was highly reliable and that plaintiffs have not offered any evidence that would discredit his testimony. Likewise, the Court finds that the plaintiffs have not demonstrated that the data supporting the perpetrator analysis was flawed in any manner. Further, Dr. Siskin testified that the daily control log was crossed-checked with actual incident reports and that no significant differences were observed and that no bias was found in the data. In sum, the Court finds that the plaintiffs' challenges to the data are speculative at best.

72. Defendant also offered Dr. Moffatt's studies to support the job-relatedness and business necessity of its aerobic capacity requirement of 42.5 mL/min/kg. After reviewing Dr. Moffatt's studies and his testimony, the Court finds that Dr. Moffatt's studies demonstrate that an aerobic capacity level of less than 45 mL/kg/min resulted in a significant decrement in upper body strength after an officer undertook a .35 mile run at a pace consistent with responding to an officer assist call. Conversely, Dr. Moffatt's studies showed that individuals with an aerobic capacity of 45 mL/kg/min or above only suffered a 5% to 10% decrement in upper body strength after a .35 mile paced run simulating an officer assist call.

73. Because an officer who arrives at the scene of an officer assist call must be prepared to engage in arduous activities, such as the apprehension of a resisting perpetrator, crowd control or combative situations, the officer must possess a sufficient energy reserve upon arrival. In light of Dr. Moffatt's study, it is plain that SEPTA officers with an aerobic capacity of less than 45 mL/kg/min would be less able to engage in combative situations after the officer has engaged in a paced run than those officers who possess an aerobic capacity of 45 mL/kg/min or greater.

74. Consequently, the Court finds that Dr. Moffatt's studies demonstrate the manifest relationship of aerobic capacity

to the critical and important duties of a SEPTA transit police officer, *i.e.*, the ability to provide officer assistance on foot in critical and potentially life-threatening situations.

75. In summary, the Court concludes that the overwhelming empirical evidence demonstrates that meeting or exceeding SEPTA's aerobic capacity standard translates into increased levels of Part I arrests, increased Part I arrest rates and generally a higher proficiency for critical tasks such as pursuits, officer backups and officer assists.

76. The Court is impressed with the convergence of evidence that the commendation studies, award studies, perpetrator studies and arrest studies have in demonstrating the predictive and useful relationship between SEPTA's aerobic capacity requirement and increasing levels of arrest performance on the job. Based on the evidence admitted at trial, the Court finds that aerobic capacity predicts and correlates with arrests, which is a critical and important task of SEPTA transit police officers. Indisputably, SEPTA's aerobic capacity requirement bears a manifest relationship to the position of a SEPTA transit police officer. Therefore, SEPTA has met its burden of establishing the job relatedness and business necessity of its aerobic capacity standard.[19]

77. Before considering whether plaintiffs have established that alternative selection devices exist which would equally serve SEPTA's business goal of having a police officer workforce capable of performing the physical requirements of the job and that such alternatives would have either no adverse impact against female applicants or less adverse impact than SEPTA's physical fitness test at issue in this case, the Court will address plaintiffs' broader-based arguments attacking the business necessity and job-relatedness of the 1.5 mile run of SEPTA's physical fitness test.

78. Throughout the course of the trial, plaintiffs argued that because some incumbents have occasionally failed

---

19. With respect to the administrations of SEPTA's physical fitness test prior to November 21, 1991, the effective date of the Civil Rights Act of 1991, the Court concludes that the United States has failed to demonstrate that the 12 minute, 1.5 mile run component of SEPTA's physical fitness is not job-related or not consistent with business necessity.

SEPTA's aerobic capacity test or some aspect of the muscular strength and endurance test, SEPTA's physical fitness standards must be invalid. Essentially, plaintiffs argue that an employer can never raise standards through its applicant testing if, in fact, some incumbents are unable to achieve those standards. Even leaving aside the collective bargaining agreement issue, this Court will not accept the proposition that employers are restricted from raising standards and that they are bound in their hiring by the level of performance of its incumbent work force.

79. In 1991, it was SEPTA's mission to improve the physical ability level of its force to combat crime more effectively. SEPTA management had observed that its Transit Police Department was not effectively preventing or combatting crime due in part to its officers' low level of physical fitness. Thus, SEPTA decided to increase the fitness of its workforce by implementing applicant and incumbent physical fitness testing. The Court finds SEPTA's goal laudable and appropriate given the evidence of the high crime rate in the SEPTA system in the late 1980s and early 1990s; indeed, employers such as SEPTA should be encouraged to improve the efficiency of its workforce, especially where public safety is implicated by the particular job as it is with SEPTA. Thus, if employers wish to improve the effectiveness and efficiency of their incumbent workforce, these employers cannot be bound by the performance of their incumbent employees.

80. Dr. Zedeck, plaintiffs' expert, agreed that there are many levels of performance in a work force that range from poor to outstanding and that an employer may use an applicant test to enhance performance above simply satisfactory. Indeed, a valid test can set a cutoff score above the average performer or even above the highest incumbent performer. Dr. Zedeck agreed that SEPTA's inability to enforce its incumbent fitness program did not invalidate its aerobic capacity test. SEPTA's expert, Dr. Henderson, also testified that incumbent failures are irrelevant to the validity of a test developed as a selection device.

81. The Court finds plaintiffs' argument concerning incumbent failures wholly unpersuasive. The logical absurdity

of this argument is that no employer could ever raise standards without firing its entire incumbent work force. There exist a myriad of reasons why an employer may retain incumbents while using selection devices to raise the standards of performance of recently hired employees. The Court finds that in this case, the dramatic change in the aerobic capacity of SEPTA's transit officers is an example of how performance may be raised through applicant testing.[20] Consequently, the Court rejects plaintiffs' incumbent officer argument.

82. Plaintiffs further argue that SEPTA's aerobic capacity requirement has no job-relatedness and is not consistent with business necessity because SEPTA is not unique as a foot-based force. In essence, plaintiffs seem to argue that SEPTA cannot use its aerobic capacity requirement to select applicants because other organizations, which allegedly are similar in terms of job responsibilities, do not have such a requirement. This argument misses the mark and has no relevance to the job-relatedness and business necessity of SEPTA's aerobic capacity requirement. SEPTA's aerobic capacity requirement cannot be said to be lacking in job-relatedness or business necessity simply because other law enforcement agencies fail to use such an aerobic capacity requirement. If the Court were to credit plaintiffs' argument, then no employer could ever use a selection device that was greater than or different than those selection devices being used by other like employers. This result is not required by Title VII and would, in application, prevent employers from improving the performance of its workforce. In addition, plaintiffs have not offered any evidence establishing that these other law enforcement agencies are performing better than or even as well as SEPTA's police force. Thus, the Court rejects this argument.

83. Plaintiffs further contend that SEPTA's aerobic capacity requirement is neither job-related nor consistent with business necessity because the Philadelphia Police Department responds to "a substantial portion of the crime on the SEPTA

20. To the credit of SEPTA's transit police officers, despite the inability of the department to discipline incumbents who fail to meet their goals or standards, 84% of SEPTA's pre-physical testing hires have met SEPTA's standards.

subway and elevated system (20-36%)" and because the Philadelphia Police Department handles more crime than SEPTA does on a daily basis. The Court, however, is hard-pressed to understand how these statistics demonstrate the lack of job-relatedness and business necessity of SEPTA's aerobic capacity requirement. The Philadelphia Police Department cannot be said to have a force that performs better than or as well as SEPTA's force merely because they respond to crime on SEPTA's property and handle more crime than SEPTA on a daily basis, especially in light of the fact that the Philadelphia Police Department is substantially larger than SEPTA in terms of the number of officers employed by these law enforcement agencies - SEPTA employs approximately 300 police officers as compared to the 5,800 officers employed by the Philadelphia Police Department (nearly twenty times the size of SEPTA). Thus, the Court rejects this argument as well.

C. Alternative Selection Devices

84. Having found that SEPTA's physical fitness test is job-related and consistent with business necessity, the burden shifts to the plaintiffs (or remains, as would be the case prior to November 26, 1991) to establish that alternative selection devices exist which would equally serve SEPTA's business goal of having a police officer workforce capable of performing the physical requirements of the job and that such alternatives would have either no adverse impact against female applicants or less adverse impact than SEPTA's physical fitness test at issue in this case.

85. As an initial matter, the Court finds that Dr. McArdle's proposed alternative test, which has different absolute standards for men and women, is not prohibited by Section 106 of the Civil Rights Act of 1991 which provides:

> (l) It shall be an unlawful employment practice for a respondent, in connection with the selection of referral of applicants or candidates for employment or promotion, to adjust the scores of, use different cutoff scores for, or otherwise alter the results of, employment related tests on the basis of race, color, religion, sex or national origin.

42 U.S.C. § 2000e-2(l).

86. The Court concludes that Dr. McArdle's proposed test does not apply different cutoff scores on the basis of gender within the meaning of Section 106. Rather, the test applies the same cutoff scores in terms of requiring the same level of relative fitness for every candidate. Once that fitness level is determined, the scores are not adjusted or altered in any way.

87. Section 106 "intends only to ban the discriminatory adjustment of test scores or cutoffs." 137 Cong. Rec. H9547 (daily ed. Nov. 7, 1991) (statement of Rep. Hyde); *see also id.* ("race norming or any other discriminatory adjustment of scores or cutoff points of any employment related test is illegal"); *id.* at S15476 (daily ed. Oct. 30, 1991) (statement of Sen. Dole). Thus, Section 106 was designed to prevent the arbitrary alteration of test scores or the use of different cutoff scores based on nothing more than the fact that certain groups do not score as well on a test.

88. The physical fitness test recommended by Dr. McArdle, in contrast, neither "adjusts" scores nor applies different cutoffs solely because certain groups do not score as well on the test. Rather, the proposed test takes into account immutable physiological characteristics widely recognized in the scientific community, and uses those characteristics to evaluate each candidate's relative physical fitness - relative to other members of the same gender. When Congress enacted Section 106, it indicated that "[a]pplicants and workers of all races, ethnic groups, and genders have the right to a level playing field and to selection based on merit." 137 Cong. Rec. H9529 (daily ed. Nov. 7, 1991) (statement of Rep. Edwards).

89. The only case in which this specific issue has been decided, *Powell v. Reno*, Civil Action No. 96-2743 (D.C. July 24, 1997), rejected SEPTA's precise argument. In *Powell*, the plaintiff challenged his termination from the FBI, which was based on his failure to pass the physical fitness requirements of the FBI's training Academy at Quantico, Virginia. The Academy had different passing scores for men and women. The plaintiff alleged that these different passing scores violated Title VII by discriminating against him on the basis of his sex. In sustaining the use of different physical fitness measures for males and females, the court stated:

> Title VII allows employers to make distinctions based on undeniable physical differences between men and women.
>
> . . . .
>
> Basic physiological differences, such as discrepancies in upper body strength and size, result in males and females of similar fitness levels performing differently on physical fitness tests. Comparing men against men and women against women, the FBI's physical fitness standards appropriately take these differences into account. Accordingly, the requirements for males co-exist with comparable requirements for females.

Powell, slip op., at 6-7 (citation omitted).

90. *Powell* is consistent with other cases in which courts have upheld overall fitness requirements that contain gender differences. *See Gerdom v. Continental Airlines, Inc.*, 692 F.2d 602, 606 (9th Cir. 1982) (en banc) (collecting cases) ("[s]everal courts have similarly upheld physiologically based policies which set a higher maximum weight for men than for women of the same height"). Different weight requirements are valid as long as "no significantly greater burden of compliance was imposed on either sex; that is the key consideration." *See id.* (citations omitted); *see also United States v. City of Wichita Falls*, 704 F. Supp. 709, 714, 715 n.4 (N.D. Tex. 1988)

91. Accordingly, this Court concludes that Section 106 of the Civil Rights Act of 1991 does not prohibit the use of different passing scores on a physical fitness test based on the well-established, immutable physiological differences between men and women and that therefore Dr. McArdle's proposed alternative test does not violate Section 106.

92. Although the Court finds that Dr. McArdle's proposed alternative test does not violate Section 106, the Court finds that this alternative would not equally serve SEPTA's business goal of having a police officer workforce capable of performing the physical requirements of the job as well as its existing test does. Thus, the Court rejects Dr. McArdle's proposed alternative.

93. Under Dr. McArdle's proposed alternative, the applicants would be required to meet the normative standards pro-

posed by the Cooper Institute - relative standards of fitness for women and men. However, no evidence was presented by plaintiffs that the normative standards of the Cooper Institute are predictive of or correlate with good police officer performance. For that matter, plaintiffs' experts readily admit that there is no data to demonstrate that these normative standards correlate with any occupation, let alone law enforcement work.

94. Plaintiffs' own expert, Dr. Zedeck, flatly refused to endorse a proposed alternative test that was not validated for the SEPTA transit police officer work. Moreover, Dr. McArdle admitted that the normative standards of the Cooper Institute were not validated for police work.

95. Because plaintiffs cannot establish that the normative standards of the Cooper Institute can predict or are even correlated with successful performance as a SEPTA transit police officer, the Court cannot find that Dr. McArdle's proposed alternative would equally serve SEPTA's business goal of having a police officer workforce capable of performing the physical requirements of the job as well as its existing test does.

96. The Court also notes that the only evidence offered as to the predictive or correlative nature of Dr. McArdle's relative fitness standards showed that such relative fitness standards did not predict or correlate with good police work. In this regard, Dr. Siskin undertook a series of studies that tested whether relative fitness standards would predict performance in the various arrest parameters that he studied for SEPTA. Dr. Siskin concluded that the relative fitness model failed to predict patrol officer performance; instead, this relative fitness model showed a negative gender effect rather than a positive prediction. The conclusion that Dr. Siskin drew was that absolute aerobic capacity predicted SEPTA transit patrol officer performance, whereas relative fitness did not. Consequently, Dr. McArdle's test cannot be found to be as equally effective as SEPTA's existing aerobic capacity requirement that has been shown to be predictive of successful performance on the job as a SEPTA transit patrol officer.

97. During the course of the trial, the plaintiffs presented evidence regarding physical fitness tests from other transit authorities and police jurisdictions and argued that these tests, which have lower standards than SEPTA's test, should be adopted by SEPTA. However, according to plaintiffs' own expert Dr. Zedeck, the *Uniform Guidelines* prohibit the transportation of test from one jurisdiction to another that has not been validated, especially where there has been no demonstration through a competent job analysis that the positions are the same or substantially similar. Indeed, when confronted with Dr. McArdle's proposal, Dr. Zedeck flatly refused to endorse the transportability of invalidated tests, such as the normative standards of the Cooper Institute contained in Dr. McArdle's proposal.

98. The Court also finds plaintiffs' other proposed alternative selection devices to be unacceptable. Indeed, plaintiffs' other selection devices are actually not tests at all. What plaintiffs propose is that SEPTA simply hire people without any physical fitness testing, despite the fact that plaintiffs would even concede that some level of physical fitness is needed to be a SEPTA transit police officer, and then send these applicants to physical fitness training with the hope that they would pass this training. In essence, plaintiffs want to replace SEPTA's successful physical abilities test with no test and a risk that its untested applicants may fail the training that was paid for by SEPTA with no guarantee that any of these persons would succeed at the training. Although this proposal is patently absurd on its face, plaintiffs were able to produce many examples of such a test that was actually being used by other law enforcement agencies. However, a close review of these tests demonstrates that these tests, if they should even be called tests, would not equally serve SEPTA's business goal of having a police officer workforce capable of performing the physical requirements of the job as well as SEPTA's existing test does.

99. The Court, after considering the testimony of Chief McDevitt from the Washington D.C. Metropolitan Area Transit Authority ("WMATA"), Robin Zarbo from the AMTRAK Police Department and Chief Inspector Pryor and Lieutenant

O'Donell from the City of Philadelphia Police Department, concludes that these other law enforcement agencies, unlike SEPTA, show a disregard for the level of physical fitness of their officers by not administering any applicant physical abilities tests. Furthermore, not one of these other agencies has a validated physical abilities test. In essence, none of these agencies has conducted a study to determine whether its selection device actually predicts or correlates with successful performance as a law enforcement officer. In response to this fact, plaintiffs circularly argue that these agencies do not have to validate their tests because these tests do not have an adverse impact. This response, however, begs the question as to whether a selection device actually correlates with or is predictive of successful performance on the job. Thus, it is irrelevant that these agencies do not have to validate their tests because they do not have a large adverse impact.[21]

100. Upon review of these other agencies' testing devices, it is clear that none of these tests would equally serve SEPTA's goals as well as SEPTA's current physical fitness test. In particular, AMTRAK specifically disclaims the use of any physical abilities testing and is willing to accept applicants who fail the Academy's low level of training. With respect to WMATA, there is no physical abilities testing of applicants. WMATA's only physical abilities testing is voluntary and for the promotion of officers. Moreover, WMATA's test only requires officers to reach the thirtieth percentile under the normative standards of the Cooper Institute in order to be considered for employment with WMATA. In contrast, the Court finds that due to the physically demanding job of a SEPTA transit police officers, any use of the thirtieth percentile of the normative

---

21. The fact that many law enforcement agencies have adopted selection devices, which have not been correlated with successful job performance, in order to avoid an adverse impact on women is not surprising. Professor Cox, in *Employment Discrimination*, notes that many employers will choose to adopt non-predictive but neutral selection devices in order to avoid expensive litigation under the disparate impact theory. *See* Cox, *supra*, at 8-1 - 8-101. Thus, in order to avoid vexing and expensive litigation, employers are adopting non-predictive tests with no adverse impact, even though these employers may not necessarily be selecting the most-qualified persons for the job - a result which was never intended by Title VII.

standards of the Cooper Institute would be highly inappropriate, if not dangerously irresponsible.

101. In addition, the Philadelphia Police Department, like AMTRAK, has no physical abilities testing for applicants to its police force; instead, the Philadelphia Police Department relies on training at the Academy to prepare its hires for the job of a Philadelphia police officer. However, Lt. O'Donnell, chief trainer at the Academy, openly confessed that the 26 hours of training administered during the Academy is unsatisfactory for recruits of the Academy to obtain an acceptable level of physical fitness. Having observed the lack of utility of Academy training, Lt. O'Donnell encourages officers to train on their own outside of the Academy and has actually discussed with his supervisors the need for a validated applicant physical abilities test.

102. The Court also notes that the Academy uses a weighted scoring system based on the normative standards of the Cooper Institute that enables officers to completely fail certain components of the Academy fitness test if they do well on other components. The significance of the weighted test could have a particularly detrimental effect to SEPTA in that the aerobic capacity test - a 1.5 mile run - at the Academy could be completely diluted by a recruit's ability to do well in the non- aerobic portions of the test. Counsel for SEPTA demonstrated that applicants could complete the 1.5 mile run in no specified time as long as they are able to pass the other aspects of the Academy's test. Under plaintiffs' proposed alternative, an applicant to SEPTA could actually be hired as a patrol officer even if the applicant had low aerobic capacity. In light of the evidence that establishes that aerobic capacity is a physical ability that SEPTA patrol officers need in order to perform their duties successfully, the Court simply cannot find that plaintiffs' proposed alternative of sending applicants to the Academy for training is an alternative selection device which would equally serve SEPTA's goals.

103. The Court also finds that plaintiffs cannot establish that their proposed alternative tests will equally serve SEPTA's goals because plaintiffs have not shown job similarity between the other law enforcement agencies from whence

plaintiffs' alternative tests come and SEPTA. Unlike other law enforcement agencies, SEPTA officers patrol alone, spend a vast majority of their time on foot and engage in foot chases, stair climbing, physical arrests and an array of other physical tasks without the assistance of motorized transportation or a partner to supply backup. Therefore, physical fitness tests that are appropriate for car-based police forces may be inappropriate for SEPTA. Indeed, plaintiffs have not offered any evidence indicating that their proposed alternative tests, which are geared for car-based patrol forces, are appropriate for SEPTA.

104. In sum, the Court flatly rejects plaintiffs' proposed alternative selection devices as an alternative to SEPTA's aerobic capacity test. Unlike the other transit authorities and the Philadelphia Police Department, SEPTA already has a validated test in place which relates to the specific tasks to be performed by its officers. The Court thus will not accept the use of invalidated tests from dissimilar law enforcement agencies for use at SEPTA.

D. Gym-Based Components of SEPTA's Physical Fitness Test

105. The United States has also challenged the gym-based muscular strength and endurance test that was administered to applicants in 1991 and 1993. The Court, however, will not determine whether these tests violate Title VII because the United States challenge to these tests is now moot.

106. The applicant gym-based muscular strength and endurance test was discontinued in 1994 in favor of a criterion-based test, which the United States does not challenge here. Chief Evans has testified that SEPTA will not reimpose testing on the gym-based components of the physical fitness test. Thus, assuming *arguendo* that the United States could prevail on its challenge to the gym- based test, there simply is no present harm to enjoin. *See, e.g., Roe v. City of New Orleans*, 766 F. Supp. 1443, 1453 (E.D. La. 1991). In addition, the United States has not identified one female applicant who would be entitled to damages. Thus, the Court finds that the United States' challenge to the gym-based components of SEPTA's former physical fitness test is moot.

E. Conclusion

107. Based on the foregoing Findings of Fact and Conclusions of Law, the Court finds in favor of SEPTA and against the *Lanning* plaintiffs and the United States. Judgment will thus be entered in favor of SEPTA and against plaintiffs.

An appropriate Order follows.

Clarence C. Newcomer, J.

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF PENNSYLVANIA

CATHERINE NATSU LANNING, ET AL., *Plaintiffs,*	:	Civil Action
v.	:	
SOUTHEASTERN PENNSYLVANIA TRANSPORTATION AUTHORITY, *Defendant.*	:	No. 97-0593

UNITED STATES OF AMERICA, *Plaintiff,*	:	Civil Action
v.	:	
SOUTHEASTERN PENNSYLVANIA TRANSPORTATION AUTHORITY, *Defendant.*	:	No. 97-1161

ORDER

AND NOW, this day, of June, 1998, upon consideration of the testimony of the witnesses, the admitted exhibits, the arguments of counsel, and the parties' post-trial submissions, and consistent with the foregoing Findings of Fact and Conclusions of Law, it is hereby ORDERED that JUDGMENT is ENTERED in favor of defendant Southeastern Pennsylvania Transportation Authority and AGAINST plaintiffs Catherine Natsu Lanning, Denise Dougherty, Altovise Love, Belinda Kelly Dodson, Lynne Zirilli and the class members in Civil Action No. 97-0593. IT IS FURTHER ORDERED that JUDGMENT is ENTERED in favor of defendant Southeastern Pennsylvania Transportation Authority and against plaintiff United States of America in Civil Action No. 97-1161. The Clerk of the Court shall mark these cases CLOSED.

AND IT IS SO ORDERED.

Clarence C. Newcomer, J.

APPENDIX C

§ 2000e-2. Unlawful employment practices

(a) Employer practices

It shall be an unlawful employment practice for an employer—

(1) to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions, or privileges for employment, because of such individual's race, color, religion, sex, or national origin; or

(2) to limit, segregate, or classify his employees or applicants for employment in any way which would deprive or tend to deprive any individual of employment opportunities or otherwise adversely affect his status as an employee, because of such individual's race, color, religion, sex, or national origin.

• • • •

(k) Burden of proof in disparate impact cases

(1)(A) An unlawful employment practice based on disparate impact is established under this subchapter only if—

(i) a complaining party demonstrates that a respondent uses a particular employment practice that causes a disparate impact on the basis of race, color, religion, sex, or national origin and the respondent fails to demonstrate that the challenged practice is job related for the position in question and consistent with business necessity; or

(ii) the complaining party makes the demonstration described in subparagraph (C) with respect to an alternative employment practice and the respondent refuses to adopt such alternative employment practice.

(B)(i) With respect to demonstrating that a particular employment practice causes a disparate impact as described in subparagraph (A)(i), the complaining party shall demonstrate that each particular challenged employment practice causes a disparate impact, except that if the complaining party can demonstrate to the court that the elements of a respondent's decisionmaking process are not capable of separation for analysis, the decisionmaking process may be analyzed as one employment practice.

(ii) If the respondent demonstrates that a specific employment practice does not cause the disparate impact, the respondent shall not be required to demonstrate that such practice is required by business necessity.

(C) The demonstration referred to by subparagraph (A)(ii) shall be in accordance with the law as it existed on June 4, 1989, with respect to the concept of "alternative employment practice."

• • • •

(l) Prohibition of discriminatory use of test scores

It shall be an unlawful employment practice for a respondent, in connection with the selection or referral of applicants or candidates for employment or promotion, to adjust the scores of, use different cutoff scores for, or otherwise alter the results of, employment related tests on the basis of race, color, religion, sex, or national origin.

• • • •

§ 2000e-6. Civil actions by the Attorney General

(a) Complaint

Whenever the Attorney General has reasonable cause to believe that any person or group of persons is engaged in a pattern or practice of resistance to the full enjoyment of any of the rights secured by this subchapter, and that the pattern or practice is of such a nature and is intended to deny the full

exercise of the rights herein described, the Attorney General may bring a civil action in the appropriate district court of the United States by filing with it a complaint (1) signed by him (or in his absence the Acting Attorney General), (2) setting forth facts pertaining to such pattern or practice, and (3) requesting such relief, including an application for a permanent or temporary injunction, restraining order or other order against the person or persons responsible for such pattern or practice, as he deems necessary to insure the full enjoyment of the rights herein described.

APPENDIX D

PART 1607—UNIFORM GUIDELINES ON EMPLOYEE SELECTION PROCEDURES (1978)

(Excerpts from 29 C.F.R.)

§ 1607.3 Discrimination defined: Relationship between use of selection procedures and discrimination.

A. *Procedure having adverse impact constitutes discrimination unless justified.* The use of any selection procedure which has an adverse impact on the hiring, promotion, or other employment or membership opportunities of members of any race, sex, or ethnic group will be considered to be discriminatory and inconsistent with these guidelines, unless the procedure has been validated in accordance with these guidelines, . . .

• • • •

§ 1607.5 General standards for validity studies.

A. *Acceptable types of validity studies.* For the purposes of satisfying these guidelines, users may rely upon criterion-related validity studies, content validity studies or construct validity studies, in accordance with the standards set forth in the technical standards of these guidelines, section 14 below. New strategies for showing the validity of selection procedures will be evaluated as they become accepted by the psychological profession.

B. *Criterion-related, content, and construct validity.* Evidence of the validity of a test or other selection procedure by a criterion-related validity study should consist of empirical data demonstrating that the selection procedure is predictive of or significantly correlated with important elements of job performance. Section 14B below. Evidence of the validity of a test or other selection procedure by a content validity study should consist of data showing that the content of the selection procedure is representative of important aspects of performance on the job for which the candidates are to be evalu-

ated. See 14C below. Evidence of the validity of a test or other selection procedure through a construct validity study should consist of data showing that the procedure measures the degree to which candidates have identifiable characteristics which have been determined to be important in successful performance in the job for which the candidates are to be evaluated. See section 14D below.

C. *Guidelines are consistent with professional standards.* The provisions of these guidelines relating to validation of selection procedures are intended to be consistent with generally accepted professional standards for evaluating standardized tests and other selection procedures, such as those described in the Standards for Educational and Psychological Tests prepared by a joint committee of the American Psychological Association, the American Educational Research Association, and the National Council on Measurement in Education (American Psychological Association, Washington, DC, 1974) (hereinafter "A.P.A. Standards") and standard textbooks and journals in the field of personnel selection.

• • • •

H. *Cutoff scores.* Where cutoff scores are used, they should normally be set so as to be reasonable and consistent with normal expectations of acceptable proficiency within the work force. Where applicants are ranked on the basis of properly validated selection procedures and those applicants scoring below a higher cutoff score than appropriate in light of such expectations have little or no chance of being selected for employment, the higher cutoff score may be appropriate, but the degree of adverse impact should be considered.

APPENDIX E

Principles for the Validation and Use of Personnel Selection Procedures: Third Edition

Society for Industrial and Organizational Psychology, Inc. 1987

Principles for the Validation and Use of Personnel Selection Procedures Statement of Purpose

This statement of principles has been adopted by the Executive Committee of the Society for Industrial and Organizational Psychology (Division 14 of the American Psychological Association) as the official statement of the Society concerning procedures for conducting validation research and personnel selection. Its purpose is to specify principles of good practice in the choice, development, evaluation, and use of personnel selection procedures.

Such selection procedures include, but are not limited to, standardized paper-and-pencil tests, performance tests, work samples, personality inventories, interest inventories, projective techniques, tests of honesty or integrity including polygraph examinations, assessment center evaluations, biographical data forms or scored application blanks, interviews, educational requirements, experience requirements, reference checks, physical requirements such as height or weight, physical ability tests, appraisals of job performance, computer-based test interpretations, estimates of advancement potential, or any other selection instrument, whenever any one or a combination of them is used or assists in making a personnel decision. Personnel decisions are decisions to hire, promote, terminate, transfer and/or take other actions that affect one's employment status.

Selection procedures are used with the implicit assumption that some important aspect of behavior on the job (including performance in training, advancement, termination, or other organizationally pertinent behavior as well as quality

or quantity of job performance) can be predicted from scores on the selection procedure. The essential principle in the evaluation of any selection procedure is that evidence be accumulated to support an inference of job relatedness.

This document is a revision of the second edition of the **Principles for the Validation and Use of Personnel Selection Procedures** published in 1980. The revision is stimulated by theoretical and research developments since the previous edition and the publication of **Standards for Educational and Psychological Testing** in 1985 by the American Educational Research Association, American Psychological Association and the National Council on Measurement in Education. The purposes of the revision are to bring the **Principles** up to date scientifically, to make them consistent with the **Standards,** and to reduce possible ambiguities regarding good practice in the use of selection procedures in making employment decisions.

B. Use of Research Results. The researcher should consider the use of the research and the reported findings by the employer. The following principles apply:

3. Cut off or other critical scores may be set as high or as low as the purposes of the organization require, if they are based on valid predictors. This implies that (a) the purposes of selection are clear and (b) they are acceptable in the social and legal contest in which the employing organization functions. In usual circumstances, the relationship between a predictor and a criterion may be assumed to be linear. Consequently, selecting from the top scorers on down is almost always the most beneficial procedure from the standpoint of an organization if there is an appropriate amount of variance in the predictor. Selection techniques developed by content-oriented procedures and discriminating adequately within the range of interest can be assumed to have a linear relationship to job behavior. Consequently, ranking on the basis of such scores is appropriate. It is not necessary to add any assumptions about underlying traits in order to rank the applicants. In some circum-

stances, such as those where a production line limits the speed at which a worker can produce, a critical score may be in order.

Judgment is necessary in setting any critical or cutoff score. A fully defensible empirical basis for setting a critical score is seldom, if ever, available. The only justification that can be demanded is that critical scores be determined on the basis of a rationale which may include such factors as estimated cost benefit ratio, number of openings and selection ratio, success ratio, social policies of the organization, or judgments as to required knowledge, skill, or ability on the job. If critical scores are used as a basis for rejecting applicants, their rationale or justification should be made known to users.

APPENDIX F

LAW OFFICES
SAUL H. KRENZEL & ASSOCIATES
SUITE 800
THE ROBINSON BUILDING
42 SOUTH 15TH STREET
PHILADELPHIA, PENNSYLVANIA 19102

SAUL H. KRENZEL
STEPHEN T. GIACCHINO*
GREGORY S. HYMAN*
SUSAN R. ENGLE
LORI KLINER-KRENZEL*
JOSHUA L. GILMAN*

*Also member of N.J. Bar

Telephone (215) 977-7230
FAX (215) 977-7240

September 23, 1999

William Suter, Clerk
Supreme Court of the United States
One First Street, N.E.
Washington, DC 20543

Re: Lanning v. SEPTA

Dear Mr. Suter:

Pursuant to our phone conversation with Loretta Ruffin, we are enclosing one copy of a video tape of the 1993 applicant run, which is marked as Appendix F to the Petition for the Writ of Certiorari, along with 40 copies of the Petition for the Writ of Certiorari to be filed with this Court. The videotape was introduced as a defense exhibit in the District Court and is discussed by the District Court in its Findings of Fact (App. B. at 80a-83a) and discussed in the dissenting opinion in the Third Circuit (App. A at p. 29a).

Very truly yours,

SAUL H. KRENZEL

SHK:mt
Enc.

cc: Lisa M. Rau, Esquire
Jules Epstein, Esquire
Michael Churchill, Esquire
Leslie A. Simon, Esquire
Robert S. Libman, Esquire